LBNL-46816

Analysis of a State Data on In-Use Emissions from Motor Vehicles (EPA Interagency Agreement DW89938103-01-1): Review of Work Completed

Tom Wenzel, Robert Sawyer

August 2000

Energy Analysis Department
Environmental Energy Technologies Division
Lawrence Berkeley National Laboratory
University of California
Berkeley, CA 94720

This work was supported by the Office of Mobile Sources, Office of Air and Radiation, of the U.S. Environmental Protection Agency. Prepared for the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

A. Introduction

This report summarizes the work completed by Lawrence Berkeley National Laboratory for EPA's Office of Transportation and Air Quality, under EPA contract number DW89938103-01-1. This interagency agreement involved the collection and analysis of data from state Inspection and Maintenance (I/M) programs. The original scope of work is attached as Appendix A. The The original contract period was June 1997 through September 1998. The agreement was extended for another year, from October 1998 through September 1999. The scope of work was modified slightly at that time. A no-cost extension was granted through June 2000.

Over the past two years LBNL has compiled and analyzed data from vehicle emissions inspection and maintenance (I/M) programs managed in several states. This effort has involved compiling four years of IM240 data from the Arizona and Colorado programs (1995 through 1998), and two years of data from the Wisconsin program (mid-1996 to mid-1998). In addition, we have obtained several years of I/M data from a state operating basic program (Minnesota), one and a half years of remote sensing data collected in the Phoenix area. As part of another project we have obtained three years (1997 through 1999) of ASM and idle data, as well as random roadside pullover tests and remote sensing data, from California. All data sets have been quality controlled and analyzed for validity and internal consistency.

State I/M data provide a rich resource on in-use vehicle emissions. The large number of vehicles tested can make up for several of the limitations of how the data were collected (inconsistent preconditioning, different test durations, etc.). This report summarizes how we used the data to accomplish the tasks set our in our original scope of work. Each section is identified by the task letter in the scope of work.

B. Evaluate the effectiveness of remote sensing devices (RSD) in accurately identifying high- and low-emitting vehicles

A clean screen program identifies vehicles suspected of having low emissions and exempts them from regularly-scheduled I/M testing. Conversely, a high-emitter identification program requires more frequent testing of suspected high emitters. EPA has published guidance to states in implementing clean screen/high-emitter identification programs. In this task we estimated the effectiveness of a hypothetical clean screen program based on remote sensing measurements, using data from Arizona. EPA has published the final report (Appendix B) on their website, and used the results in writing their guidance to states. Our analysis found that:

- A clean screen program would result in slightly larger losses of the emissions in excess of the IM240 cutpoints than a similar pilot clean screen program tested in Colorado.
- Less than one-third of the vehicles reporting for I/M testing were measured by the extensive remote sensing network utilized in the Phoenix area. This coverage rate drops to 20% if at least two remote sensing measurements are required per vehicle.
- Blanket model year exemption of the newest vehicles would be more effective than a remote sensing clean screen. A larger fraction of the fleet could be exempted, with a smaller amount of emissions in excess of the IM240 cutpoints lost (however MY92 and newer light duty trucks account for over 20% of excess NOx emissions from light duty trucks).

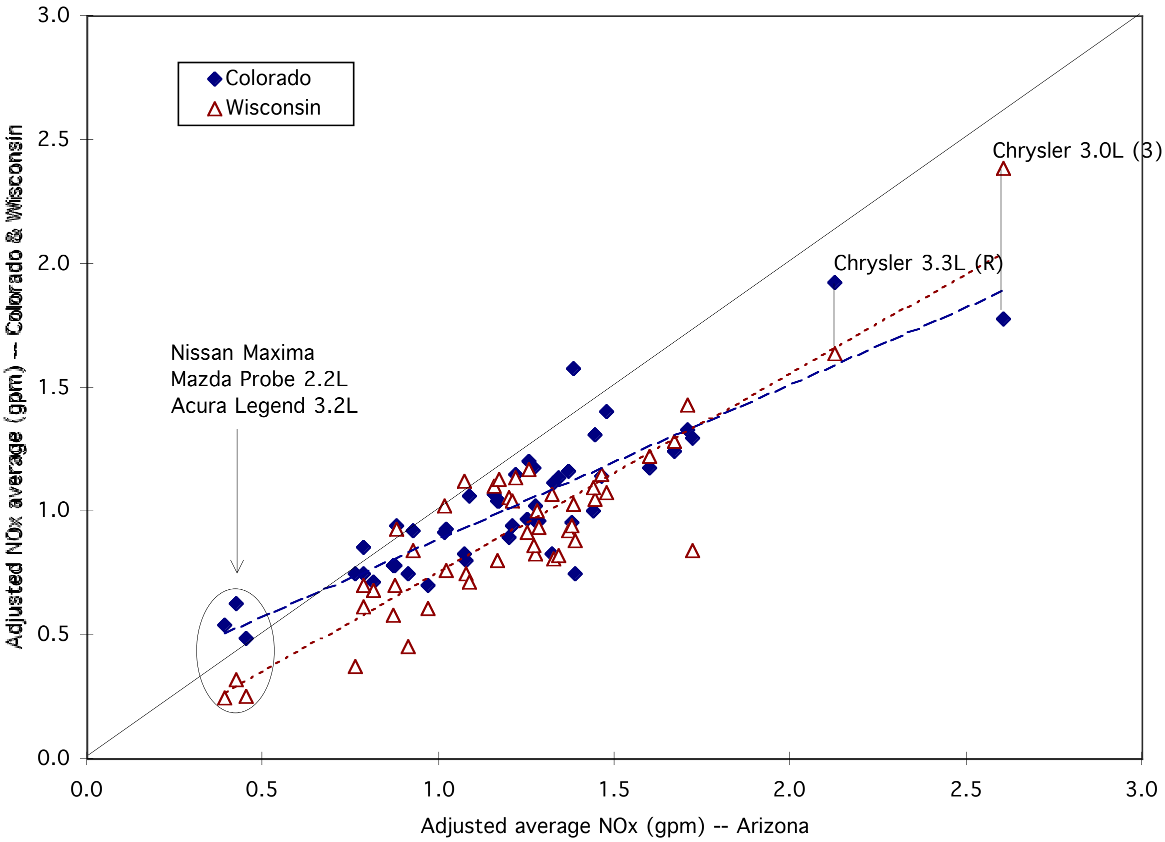
At EPA's request we also analyzed the effectiveness of using remote sensing clean screen versus model year exemptions for heavy duty (8,500 to 26,000 gvw) gasoline trucks. We found results similar to those for light duty vehicles; the results for heavy duty vehicles are summarized in a March 12, 1999 memo to Joe Somers (Appendix C).

C. Develop a list of in-use emissions by vehicle model/engine family

In this analysis we compared average emissions by vehicle model year and model, using IM240 data from three states. There is a wide range in average in-use emissions by vehicle model, suggesting that vehicle design can be improved to reduce in-use emissions. Average emissions by model are quite consistent across three state IM240 programs. **Figure 1** shows average NOx emissions of MY91 car models with at least 100 individual vehicles tested on the IM240, in three states. The vertical axis shows average NOx emissions by model in Arizona, while the vertical axis shows average emissions in Colorado (diamonds) and Wisconsin (open triangles). The figure indicates two models with consistently high NOx emissions in each state, and three models with consistently low emissions in each state. The figure demonstrates that in-use NOx emissions of the dirtiest car models can be as much as 3 times higher as those of the cleanest models of the same age.

Models tested at I/M stations in relatively low-income areas have consistently higher emissions than the same models tested at stations in relatively high-income areas. We believe this result demonstrates that in-use emissions are sensitive to vehicle maintenance, and that some vehicle models are less sensitive to maintenance practices than others. Average emissions by model using IM240 data do not correlate well with average emissions using remote sensing data, in part because remote sensing measures emissions concentrations, while the IM240 measures mass emissions (Appendix D). An adjustment to account for the fuel economy of each vehicle model, or conversion of emissions measured under each test to grams of pollutant per gallon of fuel, is needed to eproperly compare emissions across measurement techniques. As a member of the Mobile Sources Technical Review Subcommittee, Innovative and Incentive Based Transportation Policies Workgroup, we incorporated these findings into the Workgroup's recommendations to EPA. We also presented these results at CRC's On-Road Vehicle Emissions Workshops in 1997 (Appendix E), 1998 (Appendix F) and 1999 (Appendix G).

Figure 1. Average NOx by Car Model in Three States
MY91 Passenger Cars with at least 100 IM240 tests (July 1996 -- June 1997)



D. Obtain and organize in-use emissions data

As discussed above, we collected several years of IM240 data from programs in Arizona, Colorado, and Wisconsin. These databases also include idle tests of vehicles registered in Basic I/M areas of those states. In addition, we obtained an extensive database of remote sensing data from Arizona, as well as idle test data from the Minnesota Basic I/M program. We loaded these databases onto a unix workstation, and manipulate them using the SAS statistical package.

In our analysis of IM240 data from different states, we found two major limitations of comparing data across states: 1) the use of fast-pass/fast-fail algorithms complicates the comparison of emissions across vehicles, and 2) a variation in measured emissions and failure rates by season of the year. We summarize the results of our analysis of these two limitations below.

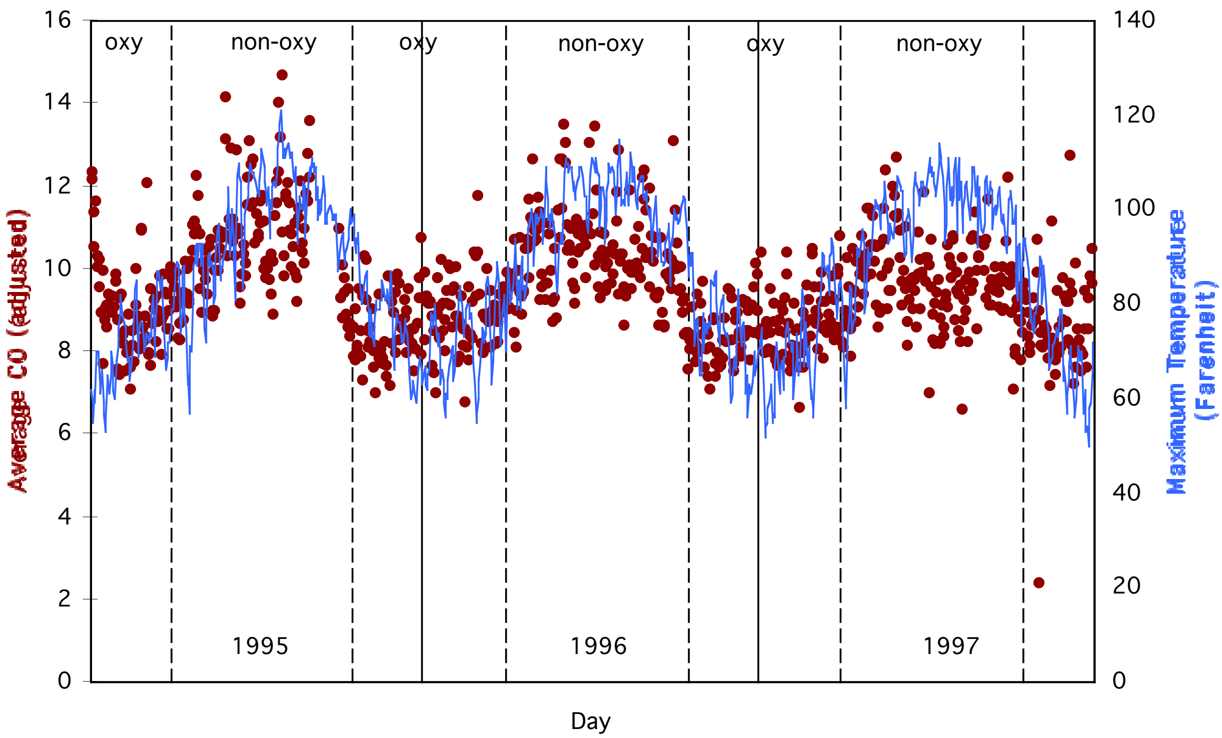
Adjusting Short Test Emissions to Full Test Equivalents

One limitation of I/M data on in-use emissions is that all vehicles are not tested over the same controlled test procedure. For instance, most states using the IM240 allow the majority of clean vehicles to pass after only 30 seconds of the 240 second test; these short tests are referred to as "fast-pass" tests. In addition, some states, such as Arizona, allow dirty vehicles to fail after only 94 seconds of testing ("fast-fail" tests). Full IM240 emissions must be estimated before comparing emissions of vehicles tested over different durations of the IM240. We developed a simple method to convert fast-pass/fast-fail emissions results to full IM240 equivalents. We compared this method to methods developed by other researchers and EPA, to get a sense of the bias introduced by using each method. We found that each method tends to underestimate full IM240 emissions from fast-passed vehicles. The findings were summarized in a June 17, 1999 memo to EPA (Appendix H).

Seasonal Variation in In-Use Emissions

We used multiple years of data from several states to examine average daily emissions from initial IM240 tests of passenger cars. We compared seasonal variation in average daily emissions with maximum daily temperature and changes in fuel composition (winter oxygenates in Arizona and Colorado, reformulated gasoline (RFG) in Wisconsin, late introduction of RFG in Arizona). There is a large seasonal variation in average emissions and I/M failure rates; **Figure 2** demonstrates how the trend in average daily IM240 CO emissions (circles) mirrors the trend in maximum daily temperature (gray lines) in Phoenix. However, these variations are not consistent across I/M programs. These results are summarized in an October 28, 1999 memo to EPA. (Appendix I) We also compared emissions by second from Arizona vehicles tested in the winter with those from vehicles tested in the summer. Seasonal trends suggest that changes in fuel composition could be a factor contributing to the changes in emissions, but more research is necessary.

Figure 2. Daily Average CO (adjusted)
Initial Tests of Passenger Cars, 1995-97 Arizona IM240



Task E. Analysis of in-use data

We performed three types of analysis under this task to demonstrate the value of using state I/M program data to analyze in-use vehicle emissions. The three types of analysis are: 1) evaluation of I/M program effectiveness; 2) analysis of in-use emissions deterioration; and 3) using in-use emissions data to model the contribution of high emitting vehicles to total fleet emissions. The results of each of these analyses is summarized below.

Evaluation of I/M Program Effectiveness

We have done an extensive evaluation of the Arizona IM240 program, using both program test results and remote sensing data. Analysis of IM240 test results from Arizona confirms EPA's findings (Glover and Brzezinski 1997) regarding the program's overall effectiveness. The initial reduction in fleet emissions due to repair compares well with TECH5 predictions for CO and HC, on the order of 15%, while the initial reduction for NOx is only half that predicted by TECH5 (7% as opposed to the predicted 17%). However, our research found that the effectiveness of I/M programs is affected by:

- Program avoidance: One-third of all vehicles that fail initial I/M testing never receive a passing test. About one-third of these vehicles are observed by remote sensors still being driven in the I/M area more than two years after their last (failing) I/M test. In addition, 40% of all vehicles tested in the first year of Enhanced program in Arizona (1995) were not tested in the next cycle (1997).
- Ineffective repair: 40% of vehicles that failed their initial test in 1995 failed again in 1997. About half of these failed for the same combination of pollutants in both years. Remote sensing data indicate that repair effectiveness drops as soon as a few months after a vehicle's final I/M test.
- Pre-test repairs: The remote sensing data also indicate that average emissions decrease substantially about three weeks prior to the initial I/M test, presumably due to pre-test maintenance, repairs, and/or adjustments (see **Figure 3**).

Analysis of vehicles reporting for two I/M cycles reveals that the effect of two years of emissions deterioration outweigh the effect of the I/M program; the after-program emissions in 1997 are substantially higher than the after program emissions in 1995 (**Figure 4** shows initial and final IM240 test emissions for the tested fleet, in tons per day, in 1995 and 1997). In addition, there are large fluctuations in the vehicle fleet tested in each I/M cycle: 40% of the vehicles tested in 1995 do not report for testing in 1997, while 40% of the vehicles tested in 1997 were not tested in 1995.

Figure 3. Average CO RSD Emissions by Time Period
1996-97 Arizona Remote Sensing

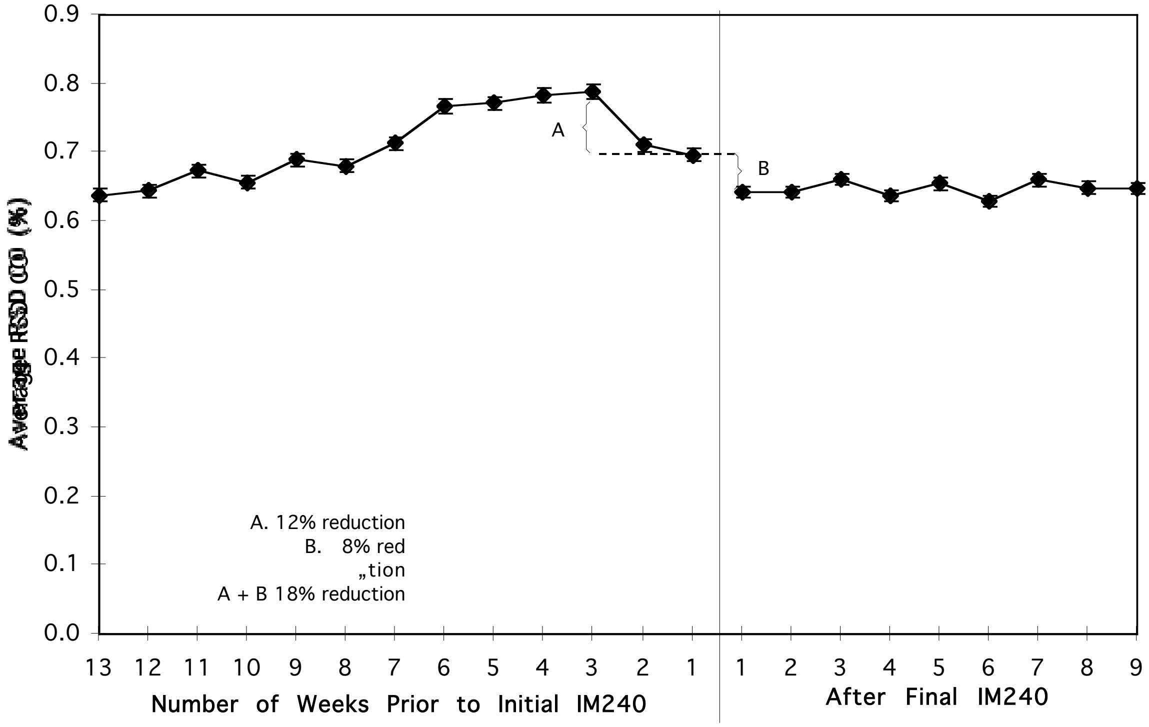
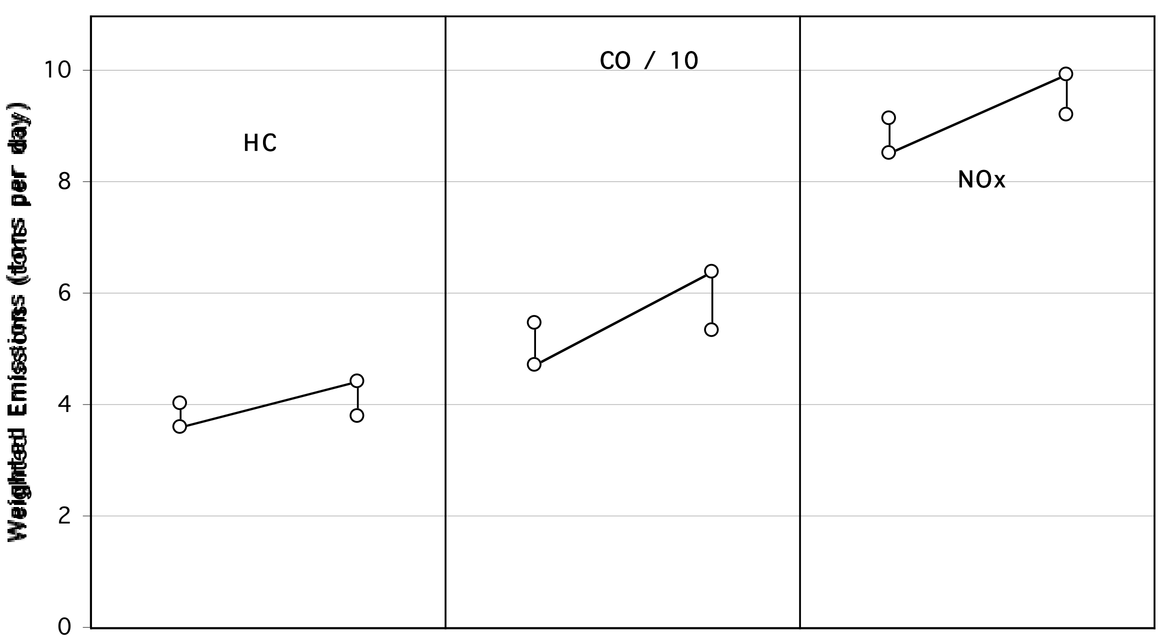


Figure 4. Fleet Emissions over Two I/M Cycles
Passenger Cars tested in both 1995 and 1997, Arizona IM240



These results are summarized in a July 27, 1999 memo to EPA (Appendix J), and have been presented at the NCVECS Clean Air Conference in 1998 and 1999 (Appendix K), to the NRC Committee to evaluate MOBILE in March 1999 (Appendix L), and at the 1999 CRC Workshop (Appendix M). These data were also used to prepare comments to EPA on MOBILE6 (Appendix N). We recently presented lessons learned from our I/M evaluation activities to a NRC Committee on Effectiveness of Vehicle Inspection and Maintenance Programs (Appendix O).

We participated in an informal group of experts assembled by EPA's Transportation and Regional Programs Division to help write guidance to states on how to evaluate I/M programs using remote sensing data. We played an instrumental role in writing the draft guidance (forthcoming). In addition, we analyzed the Arizona I/M program using the three methods of remote sensing analysis in the draft guidance, the Step, Reference, and Comprehensive methods (the comprehensive method is the one we originally applied to the Arizona program.)

At EPA's request we analyzed the emission reduction potential lost by not repairing gross emitters in the Arizona fleet. Repairing all vehicles that never pass IM240 testing (including the 4% of failed vehicles that receive waivers) would nearly double the effectiveness of the Arizona I/M program; CO and HC emission reductions would be increased from 14% to 25%. Only about half of these additional emission reductions are attributable to vehicles identified as gross emitters by at least one remote sensing measurement. We provided these results in a February 4, 1999, report to Jim Bagian (Appendix P).

We also have examined program effectiveness for all vehicles subject to I/M testing in the Phoenix area, including older vehicles subject to loaded idle testing, under a contract with the Arizona Department of Environmental Quality. These results are summarized in a report finalized in December 1999 (Appendix Q).

Under a contract with the California Inspection and Maintenance Review Committee (IMRC), we provided an in-depth evaluation of California's Enhanced Smog Check Program, Smog Check II. The evaluation made use of millions of I/M test records, 30,000 roadside tests of randomly selected vehicles, 150,000 remote sensing measurements, and nearly 50 million vehicle registration records. Because California requires additional I/M testing when vehicle ownership is changed, we were able to analyze multiple cycles of I/M test results in a 12-month time period on a portion of the vehicle fleet. Our analysis of the California multi-cycle vehicle fleet found that:

- 20% of vehicles failing their initial I/M test and passing a retest ("fail-pass" vehicles), and 6% of vehicles passing their initial I/M test ("initial pass vehicles"), failed a subsequent initial test one month later. These vehicles failed so soon after passing an I/M test either because of inherent test-to-test variability in vehicle emissions, or fradulent test practices.
- The failure rate and average emissions of the fail-pass fleet remain fairly constant over the next 9 to 12 months, indicating that any repairs made to these vehicles are effective and durable.
- The failure rate and average emissions of the initial pass fleet increase dramatically over the next 12 months, with the failure rate nearly tripling.

- Comparison of initial to final I/M test results over a single cycle of I/M testing overstates emission reductions. A more accurate estimate of emission reductions involved comparing initial to initial I/M test results over subsequent I/M cycles.

We also analyzed the degree of program avoidance, the effect of pretest repairs and maintenance, repair effectiveness by station type, and other program elements. We estimated the total tons per day exhaust emissions benefit of the Enhanced program, by source of emission reduction and vehicle model year. The IMRC approved a report of our findings at their June 19th meeting; the full report is posted on the Bureau of Automotive Repair's website (http://www.smogcheck.ca.gov/smogweb/IMRC).

In-Use Emissions Deterioration

We analyzed average emissions by model year and odometer reading to examine trends in emissions deterioration as vehicle technology improves. Analysis of average emissions of MY93 and newer vehicles at high mileage (100,000 to 200,000 miles) indicates that newer technology is more durable. Tom presented these results to the MSTRS In-Use Deterioration Workgroup in March 1997, and at the SAE Government/Industry Meeting in May 1997 (Appendix R).

Modeling High Emitters

With Marc Ross of the University of Michigan, we identified four types of high emitters (running rich, running lean, misfire, bad catalyst), using second-by-second data from the UC Riverside modal emissions model project. We determined the distribution of each type of high emitter in the in-use fleet, using Phoenix IM240 data. The findings were presented at the 1998 SAE Fuels and Lubricants meeting (981414) (Appendix S). This analysis was incorporated into the high emitter module of the Comprehensive Modal Emissions Model, or CMEM, we developed in conjunction with researchers at UC Riverside and the University of Michigan (http://www.cert.ucr.edu/groups/tsr/em.html). We have also used the IM240 data from three states to validate projected IM240 emissions from CMEM. Finally, we presented a poster at the 2000 CRC Workshop on a comparison of the incidence of high emitters, and their emissions, as measured under IM240 and remote sensing programs. We found that emissions distributions expressed as ratios to CO_2 emissions agree well between the two measurement techniques, and that the fraction of high emitters in the fleet has decreased dramatically in the 1990s, presumably due to better design of vehicle emissions controls (Appendix T).

Appendices

Appendix	Document	Date	Pages
A.	Inter-Agency Agreement Scope of Work	May 1997	4
B.	Clean Screen Report (LBNL-41918)	October, 1998	20
C.	HDT Clean Screen Memo	March 12, 1999	5
D.	Comparison of Emissions by Model Memo	July 24, 1998	4
E.	1997 CRC Presentation (LBNL-41451)	Apri1, 1997	14
F.	1998 CRC Poster	April, 1998	13
G.	1999 CRC Poster (LBNL-44157)	April, 1999	10
H.	Short Test Conversion Memo	June 17, 1999	17
I.	Seasonal Variation Memo	October 28, 1999	7
J.	Two I/M Cycles Memo	July 27, 1999	6
K.	1999 NCVECS Presentation	September 16, 1999	7
L.	1999 NRC Presentation	March 4, 1999	15
M.	1999 CRC Presentation (LBNL-44156)	April 21, 1999	8
N.	MOBILE6 Comments	July 27, 1999	12
O.	2000 NRC Presentation	February 15, 2000	14
P.	Repairing Gross Emitters Report (LBNL-44159)	February 4, 1999	8
Q.	AZ DEQ Report (LBNL-46114)	December 9, 1999	28
R.	1997 SAE Govt/Industry Presentation	May, 1997	17
S.	SAE High Emitter Type Paper (SAE 981414)	March 1998	16
T.	2000 CRC Poster	March 28, 2000	15

APPENDIX A

May, 1997

Scope of Work
Analysis of State Data on In-Use Emissions from Motor Vehicles

A. Introduction

Strict regulatory standards, and subsequent development of emission control technology, have resulted in significant reductions in tailpipe emissions from new motor vehicles; however, the corresponding regulations to ensure that cars continue to meet standards as they age and accumulate mileage may not have been nearly as effective. The Clean Air Act Amendments of 1990 charged the US EPA with developing new programs to reduce in-use emissions from motor vehicles. One new approach is allowing states to adopt different and more flexible vehicle emission inspection and maintenance (I/M) programs, as long as each state can demonstrate that its program is effective. Emissions measurements on large numbers of in-use vehicles could be extremely useful in getting a better understanding of the causes of increases in in-use emissions. And, perhaps most importantly, such information can be used either as an input for, or a check on, the MOBILE model used to estimate emissions for air quality modeling.

This project will assist EPA's Office of Mobile Sources in the collection and analysis of data generated by state I/M programs, in order to better assess in-use vehicle emissions, and to suggest new approaches to reduce them. LBNL's analysis will be divided into the following three general tasks. Examples of possible specific subtasks are listed under each general task.

B. Evaluate the effectiveness of remote sensing devices (RSD) in accurately identifying high- and low-emitting vehicles

Remote sensing device (RSD) technology has been proposed as a supplement to I/M programs; several states such as Arizona and California are using RSD, while many more are evaluating whether to use it. EPA has issued guidelines to states for determining interim regulatory emission reduction credits for implementing RSD in their I/M programs (EPA 1996). However, the analysis done to support this guidance does not:

1) Account for the use of multiple RSD readings of an individual vehicle to reduce the number of vehicles failing RSD that later pass the confirmatory I/M test (errors of commission or "false failures");
2) Provide guidance on how many RSD readings are needed for an optimal clean screening (with an acceptably low rate of errors of omission, or "false passes");
3) Utilize recent RSD and IM240 data from the real-world Arizona program; or
4) Address how the length of time between the RSD and I/M test affects the usefulness of the RSD test as a predictor of the I/M test that has already been collected by the state.

LBNL will obtain more recent RSD and I/M data. RSD and I/M test results for individual vehicles will be matched and analyzed to determine the accuracy of RSD in identifying individual, or groups of, high emitters (based on confirmatory I/M tests). The analysis will examine a variety of factors (on a model year basis), such as:

- How the selection of RSD pass/fail cutpoints affects the number of false passes and false failures, and the amount of excess emissions (over a regulatory standard) detected;
- How the averaging of multiple RSD readings within specific time frames affects the accuracy of RSD;
- What criteria (e.g., what cutpoints and how many RSD readings within what time periods) should be used in clean screening to maximize its effectiveness (e.g., assure a reasonable number of vehicles can pass the clean screening criteria yet few vehicles that pass these criteria are actually IM failures);
- How does recent RSD technology and the use of HC (or NOx) specific channels affect the RSD credits;
- How effective measurements of speed and/or acceleration, and instrument siting, are in reducing the number of false failures and passes; and
- Whether RSD data can be used to quantify the emission contributions from vehicles registered outside of, but operate within, the I/M area, or failed vehicles that never receive a follow-up I/M test (and "drop out" of the I/M system).

This analysis shall be done in a fashion suitable for EPA to utilize for MOBILE6, which will include a revision of the utility program to calculate emission credits from adoption of a RSD program.

C. Develop a list of in-use emissions by vehicle model/engine family

In a recent analysis (Wenzel and Ross 1996, Ross et al 1996) of remote sensing data on 2- to 5-year old cars, LBNL found a strong relationship between vehicle model and malfunction probability: some models have almost no malfunctions, while a few relatively inexpensive models have a malfunction probability several times that of all other models (22 percent versus 6 percent). Examination of four sets of dynamometer data (collected over FTP and IM240 cycles) confirmed that these 5 models had malfunction probabilities dramatically higher than that of all other models. Analysis of IM240 testing of over 200,000 cars in Phoenix suggests that this relationship between vehicle model and malfunction probability has continued, even after these particular models have aged as much as 9 years. The IM240 data also indicate that some new technology models (MY91-93) also exhibit high IM240 failure rates before 50,000 miles of use (Wenzel 1996). It appears that some models are more sensitive to insufficient or improper maintenance or heavy use than others; the emissions controls on these models are not as durable as those on other models. The identification of certain models with high failure rates can be used to improve the ability of current I/M programs to identify and repair malfunctioning vehicles, as well as to develop other approaches to reducing in-use vehicle emissions.

In this general task LBNL will continue to explore emission test failure rates by vehicle models/engine families. This will involve collecting in-use emissions data from other states, especially those that subject a random sample of the fleet to full IM240 testing. Likely subtasks for each set of data include:

- Evaluation of the relative quality of the data, based on checks of internal consistency and visits to representative testing facilities;
- Comparison of model/engine failure rates from a variety of measurement technologies and methods (ASM, IM240, FTP) in order to validate the relationship between failure rate and vehicle model/engine; and

- Identification of technological and social differences that may explain high failure rates of particular models/engines; technological differences can include items such as types of fuel injection (e.g., port fuel injection, throttle body fuel injection) and on-board diagnostics (OBD-1 and OBD-2).

D. Obtain and organize in-use emissions data

As more states adopt high-tech I/M testing, more data on in-use emissions will become available. In order for this data to be of value to researchers and modelers, data should be of high quality and presented in a consistent manner. In this general task LBNL will give input for standard reporting requirements for I/M data, based on its assessment of what is needed considering analysis of existing databases. LBNL will also continue to collect new data as they become available. LBNL will not only collect I/M data from states and other similarly available data but will also standardize its format making the data accessible to EPA and others. Several specific subtasks are possible, such as:

- Assessment of problems with current data collection/formatting in state I/M programs (and other sources of in-use data) and determination of what improvements could be made;
- Development of recommended vehicle testing formats to ensure that states collect the most important data, and that the data are of high quality and internally consistent;
- Development of reporting standards (such as variables to be included, sampling methodologies, method of data storage and transfer, etc.) to simplify comparison of data between states;
- Development of a process to monitor the testing of vehicles (including site visits) and reporting of data assuring data flow to EPA and others doing data anlayses is efficient and optimal;
- Make and promote recommendations for a multi-state data collection and flow process, for operation when doxens of states are conducting the “0.1 percent” random sample of mass emissions transient testing; and
- Collection and standardization of I/M-type data for further analysis by others.

E. Analysis of in-use data

LBNL will also recommend and perform analysis on the in-use data from the states. The following are the types of tasks that could be performed.

- Recommend standard types of analysis for states or EPA to perform on data once collected, perhaps even providing utility programs that operate on a database that has been put in standard form;
- Using the data from different states, LBNL may perform a variety of analyses:
  - Calculation of fleetwide and model/engine-specific emission deterioration rates by model year and mileage; and
  - Analysis of failure rates by state and/or other parameters (such as model year, technology types, even testing station) to determine what variables (such as socioeconomic factors, time in line for I/M test) have an impact on in-use vehicle emissions.

F. Reports

Reports will be produced on an annual basis, and, as appropriate, at the completion of each of the individual tasks (B, C, D, and E) listed above.

The reports will contain the following:

- a narrative of the work done
- a description of the data used (if convenient, the raw data may be included in an appendix or electronic format)
- figures and tables as appropriate
- results and conclusions.

References

Ross, M., R. Goodwin, R. Watkins, M.Q. Wang, T. Wenzel. 1995. *Real-World Emissions from Model Year 1993, 2000, and 2010 Passenger Cars*. Lawrence Berkeley National Laboratory . LBL-37977. November

U.S. EPA. 1996. "User Guide and Description for Interim Remote Sensing Program Credit Utility." September.

Wenzel, T., M. Ross. 1996. "Emissions from Modern Passenger Cars with Malfunctioning Emissions Controls," SAE Technical Paper 960067.

Wenzel, T. 1996. "In-Use Emissions by Car Model". Presentation at US EPA Office of Mobile Sources. September 10.

APPENDIX B

LBNL-41918

Analysis of a Remote Sensing Clean Screen Program in Arizona

Tom Wenzel, Robert Sawyer

October 1998

Energy Analysis Department
Environmental Energy Technologies Division
Lawrence Berkeley National Laboratory
University of California
Berkeley, CA 94720

This work was supported by the Office of Mobile Sources, Office of Air and Radiation, of the U.S. Environmental Protection Agency. Prepared for the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

Abstract

This report analyzes the effectiveness of using remote sensing data to "clean screen" vehicles for exemption from their upcoming emissions inspection and maintenance (I/M) test. By exempting the cleanest vehicles from I/M testing, limited resources can be concentrated on identifying and repairing vehicles with the highest emissions. We apply the methodology used in the Colorado clean screen pilot program to 18 months of remote sensing and IM240 data from Arizona. We analyze a random sample of vehicles given the full IM240 test, as well as all vehicles tested during this period in Arizona. Our primary conclusions are that:

1) less than one-third of the vehicles in the Arizona I/M program were measured by remote sensing over an entire year of measurement;

2) the clean screen methodology used in Colorado is slightly less effective in identifying the cleanest vehicles in the Arizona fleet;

3) exempting the newest model years of cars is more effective than using a remote sensing clean screen; however, because much of the excess NOx emissions from light duty trucks comes from relatively new model years, states should carefully consider whether to exempt recent model year trucks from I/M testing.

This report describes these results, and examines other aspects of applying a clean screen program in Arizona.

Introduction

This report summarizes progress to date on analyzing the effectiveness of using remote sensing data to identify individual vehicles that are suspected of being cleaner than the average vehicle. If the effectiveness of "clean screening" can be demonstrated, states can use remote sensing to exempt vehicles from scheduled vehicle emissions inspection and maintenance (I/M) testing, in order to reduce the costs of their programs. There are three known evaluations of pilot clean screening programs in Colorado [1] and Arizona [2, 3]. This study uses remote sensing and IM240 data collected in Arizona to evaluate the effectiveness of the Colorado pilot clean screen methodology, as described in [1], in identifying clean vehicles in the Arizona I/M fleet. In this report we first describe how we matched remote sensing measurements with IM240 measurements of the same vehicle. Then we discuss the results of applying the cutpoints used in the Colorado clean screen pilot to the Arizona data. We also examine what fraction of the IM240 fleet had sufficient remote sensing measurements to qualify for clean-screening, and compare our clean screen results with results from exempting entire model years of vehicles from I/M testing.

Methodology

In this section we discuss the sources of data used, and how we matched remote sensing measurements with IM240 measurements of the same vehicle. We used two different samples of IM240 measurements, and different ways of using remote sensing measurements, to test the sensitivity of our methodology.

Three Samples of IM240 Measurements

For the analysis we obtained 18 months of IM240 measurements from the Arizona I/M program, from January 1996 to June 1997, from the contractor for the program, Gordon-Darby Inc. The entire database consists of over 1.2 million IM240 test results. We conduct our analysis of clean screening effectiveness on all initial IM240 tests conducted over 18 months (All Test sample), as well as a subset of the initial tests of vehicles randomly selected to receive a full IM240 (Random sample). Our reasoning for analyzing the two samples separately is described below.

One measure of the effectiveness of a clean screening program is the amount of excess emissions, defined as "IM240 emissions in excess of the IM240 cutpoints", attibutable to the vehicles identified by remote sensing. To accurately calculate IM240 excess emissions, one needs full IM240 test results. Arizona allows vehicles to fast-pass or fast-fail the IM240 before the full 240-second test is completed; about half of the vehicles tested fast-fail after only 31 seconds of testing. Roughly two percent of the vehicles tested in the Arizona program were randomly selected for a full IM240 test.

For his analysis of the Colorado clean screen [1], McClintock uses a random sample of full IM240 tests. The primary focus of our analysis, then, is the Arizona 2% random sample (referred to as the "Random" sample). However, because the total number of vehicles in the sample is small, we also analyze all initial tests conducted in the 18-month period, including vehicles that fast-passed or fast-failed the IM240; we call this sample the "All Test" sample. This sample is virtually the entire fleet of vehicles subject to the I/M program.

The drawback with the All Test sample is that emissions are measured over different portions of the IM240 driving cycle. To account for different test durations, we make rather simple adjustments to the emissions of vehicles that are not tested over the full IM240 cycle. These adjustments involve dividing measured grams per measured mile driven to obtain grams per mile. The grams per mile emissions are then divided by adjustment factors that vary by pollutant and by the test duration in seconds, but not by whether or not the vehicle passed or failed the IM240 or by vehicle attributes such as vehicle type or model year [4]. The result is adjusted gram per mile emissions that simulate the emissions of a given vehicle if it were run on the full IM240 cycle. The adjusted emissions are a rough approximation of the full test-equivalent emissions for an individual vehicle; the adjustments appear to be more accurate for vehicles tested over longer durations of the IM240 than vehicles passed after 31 seconds. Since Arizona does not fail high emitters until at least second 94 of the IM240, we believe the adjustment is better for the failing vehicles.

The use of adjusted emissions for vehicles not tested over the full cycle affects the calculations of excess emissions lost or retained by the clean screen, as well as the determination of whether or not a particular vehicle passes or fails the final IM240 cutpoints. Since the excess emissions only come from vehicles failing the IM240, and the adjustments are more accurate for failing vehicles, we believe using adjusted fast-pass/fast-fail test results does not introduce too much bias in the analysis. And as others have shown, inconsistent preconditioning results in many vehicles being falsely failed under Arizona's final cutpoints [5]. This bias exists in both the All Test sample and the Random sample of full test vehicles.

Remote Sensing Measurements

We also obtained over 4 million individual remote sensing readings over the same 18-month period from the remote sensing program contractor, Hughes. Two evaluations of the Hughes remote sensing equipment involving side-by-side comparison with similar instruments developed by others have found several problems with the data generated by the Hughes sensors. In particular, the studies found that:

- on average, the Hughes instrument measured both CO and HC emissions higher than measured by the other instruments;

- problems with the license plate recognition system resulted in the Hughes instrument matching vehicle license plates with remote sensing readings of a different vehicle; and

- the accuracy of the speed and acceleration measurements of the Hughes instrument is inconsistent [6].

Even with these limitations with the Arizona remote sensing data, we treat all reported measurements as accurate for this initial analysis. We hope to critically evaluate data from particular instrumented vans and sites to obtain a subset of the Arizona remote sensing data which minimizes these limitations, in a later analysis.

Because remote sensing measurements pick up the emissions variability of individual vehicles, analysts frequently require multiple measurements of individual vehicles, and often average multiple measurements. For his analysis of the Colorado clean screen [1], McClintock required that each of the last two remote sensing readings for an individual vehicle exceed both the HC and CO cutpoints; we apply similar criteria here. We discuss the sensitivity of our results to how we used the remote sensing data later in this report.

It is quite possible that vehicle owners make changes to their vehicles between the last remote sensing reading and the IM240 test, and that these changes affect vehicle emissions. If this practice is common, the ability of remote sensing readings to predict IM240 results will be reduced. McClintock restricted the remote sensing readings used to those taken within 365 days prior to the IM240 test; for this analysis we do the same. Shortening the time period between the two tests may improve the accuracy of the clean screen. However, it will also reduce the number of vehicles with useable remote sensing readings. We discuss this issue later in this report.

The McClintock analysis is based on remote sensing measurements taken at 6 sites; Hughes used over 100 remote sensing sites in the Phoenix area, in part to obtain remote sensing readings on as large a portion of the vehicle fleet as possible. Several of the Arizona sites have negative grades. In addition, starting in October 1996, Hughes measured vehicle speed and acceleration at every site; about half of the vehicles with speed and acceleration measurements were decelerating as they passed the remote sensor. CO, and to a lesser extend HC, emissions can increase dramatically under moderate to high loads, encountered when vehicles accelerate at moderate to high speeds; HC emissions can also increase dramatically during decelerations. It may be possible to improve the Arizona clean screen accuracy by using remote sensing measurements from only certain sites that are deemed efficient in collecting accurate readings, or by applying speed and/or acceleration criteria (to eliminate individual remote sensing readings of vehicles under deceleration, or moderate to high acceleration). However, we have not yet examined in detail the effect of placing these kinds of restrictions on the remote sensing readings used.

Analysis

In this section we discuss our analysis of applying the Colorado pilot clean screen program cutpoints and methodology to Arizona data. We first discuss the fraction of the Arizona I/M fleet for which useable remote sensing readings are available. Then we compare the effectiveness of the clean screen program applied to vehicles in Colorado and Arizona. Next we compare the effectiveness of the clean screen program to exempting entire model years of vehicles in Arizona. Finally, we discuss the sensitivity of our results to how we use the remote sensing data.

RSD Coverage

An important aspect of the effectiveness of a particular clean screen program is the fraction of vehicles measured by remote sensing. Even if the remote sensors are quite accurate in predicting the results of an IM240 test for individual vehicles, if only a small fraction of the vehicle fleet is measured by the sensors the effectiveness of clean screening will be reduced.[1] Previous evaluations of remote sensing have demonstrated a wide range in vehicle coverage, from 47%

1. One could argue that remote sensors focus on vehicles that contribute disproportionately to the emissions inventory, since the sensors are more likely to measure emissions from vehicles that are driven frequently. However, I/M programs currently treat individual vehicles equally, regardless of how many miles they are driven each year.

for Sacramento, a large, rather low-density urban area, to 72% for Greeley, Colorado, a relatively small community of 69,000 [7, 8].[2]

Table 1 presents the fraction of initial IM240 tests in Arizona successfully matched with at least one remote sensing reading. The table calculates the coverage obtained for each of the two samples of the IM240 data described above. About 15% of the All Test sample are retests of vehicles which failed their initial test; we excluded these retests from the analysis (the database we used to extract the Random sample codes retests differently, and we could not easily calculate this number for the Random sample). Next, vehicle records without license plate information are excluded. About 15% of all tests have license plates coded as "NP", "PP", or "OS"; these codes stand for no plate, paper or temporary plate (typically a car dealer), and out of state plate, respectively. Finally, vehicles with inaccurate vehicle identification numbers (VINs) and subsequent tests coded as initial tests are excluded.[3] The result is that 70% to 80% of the entire IM240 sample are valid initial tests.

The next panel of Table 1 shows the fraction of remote sensing readings that are valid for use in the clean screen. The use of only remote sensing readings taken within the last year reduces the number of matched readings by about one-half. The ratio of matched remote sensing readings to matched initial IM240 tests is the same in each sample, 2.95. The final panel shows the overall match rate for each of the IM240 samples. Only 31% of the vehicles in the Random sample (4,649) could be matched with at least one remote sensing reading; this fraction drops to 19% (2,914) if two readings are required. The match rate for the All Test sample is comparable to that of the Random sample.

The match rates for each of these cases may be underestimated, however, if one considers that not a full year of remote sensing testing is available for all of the vehicles in the IM240 samples. For example, a vehicle tested on the IM240 in January 1996 (the first year of the IM240 data used in this study) would have at most only one month of remote sensing readings available for a possible match, since the first month of remote sensing data used in this study is also January 1996. To determine the effect of having a whole year of remote sensing testing available for all vehicles, we calculated a match rate for vehicles in the All Test sample receiving initial IM240 tests between January 1997 and June 1997. Each of these vehicles had at least one year of remote sensing testing available. If an entire year of remote sensing testing is available, the match rate increases to 40% of the Full Test sample with at least one remote sensing reading, and 28% if two remote sensing readings are required.[4]

2. These coverages were achieved by requiring only one remote sensing measurement per vehicle; the Greeley coverage is reduced to 45% if two measurements are required for each vehicle.

3. There are several reasons why a vehicle may have multiple initial tests within a two-year period: vehicles for sale by dealers that are not fleet-licensed must be tested every 90 days; subsequent tests of vehicles that were not passed within 5 months of the initial test are coded as initial tests; some repeat initial tests are for research purposes only; a small number of audit vehicles are covertly run through the system periodically; and a prospective buyer may voluntarily test a vehicle prior to purchase (personal communication with Frank Cox, Arizona Department of Environmental Quality).

4. Part of the increase in the rate is due to relaxing the requirement of both a valid CO and a valid HC remote sensing reading. Because the Hughes instruments used in Arizona frequently gave a valid reading for CO only, we removed this requirement to determine the maximum coverage possible in the Phoenix area.

Table 1. Remote Sensing/IM240 Match Rates for Two Arizona Test Samples *(at least one RSD within 365 days)*

	Random 2%		All Tests	
	Number	%	Number	%
Useable IM240 Tests				
Total tests			1,265,867	
Less retests			*-192,230*	15%
Initial tests	18,175		1,073,637	
Less bad plates	*-3,052*	16%	*-173,494*	14%
Good plates	15,123		900,143	
Less bad VINs	*-36*	0%	*-3,383*	0%
Good VINs	15,087		896,760	
Less subsequent initial tests	*-25*	0%	*-41,726*	3%
First Initial Test	15,062	81%	855,034	68%
Useable RSD Readings				
Matches with RSD	30,988		1,760,111	
Less Readings not within 365 days	*-17,279*	56%	*-981,567*	56%
Matched Readings	13,709	44%	778,544	44%
Matched Vehicles	4,649		264,204	
Avg Numb of Rdgs per Vehicle	2.95		2.95	
Match Rate				
Valid IM240 Tests	15,062		855,034	
Matched w/at least 1 valid RSD	4,649		264,204	
Match Rate	31%		31%	
Matched w/at least 2 valid RSD	2,914		168,074	
Match Rate	19%		20%	

There is another reason why these match rates may understate the coverage rate of a clean screen program. The Arizona remote sensing program was established to identify high emitting vehicles for mandatory additional I/M testing. It is possible that drivers intentionally avoided locations where remote sensing vans were making measurements, in order to avoid the possibility of being called in for an additional I/M test. A clean screen program, which provides an incentive for drivers to intentionally drive by the remote sensors in the hope of being excused from I/M testing, would likely have higher fleet coverage.

Clean Screen Results

There are two ways to evaluate the effectiveness of a clean screen program: the false pass rate and the fraction of excess emissions retained by the program. The false pass rate is the fraction of all vehicles that pass the remote sensing screen but fail subsequent IM240 testing. In an earlier, preliminary analysis, EPA used an alternative measure of the false pass rate: the fraction of vehicles that fail the IM240 but falsely pass the remote sensing screen [2]. Since the denominator of this value (the number of IM240 failures) is smaller than the denominator of the overall false pass rate (all tested vehicles), the alternative false pass rate typically is dramatically higher.

It is useful to examine false pass rates because they are not affected by short tests in programs that utilize fast-pass/fast-fail algorithms (such as Arizona's).[5] Because different vehicles are tested over different portions of the IM240 in Arizona, emissions values are not necessarily comparable between vehicles. And the Arizona contractor has demonstrated that inconsistent

5. Although the algorithms used may falsely fast-pass individual vehicles that would have failed a full IM240 (and vice versa).

preconditioning of vehicles prior to IM240 testing can cause vehicles to be improperly failed [5]; inadequate preconditioning tends to increase emissions measurements. The drawback of analyzing false pass rates in isolation is that they treat all failing vehicles equally, without accounting for the relative emission levels of high emitting vehicles. And false pass rates become problematic when determining whether a particular vehicle would fail under the stricter final IM240 cutpoints; since this determination can only be made with the emissions measurements already made, there is the potential for certain vehicles to improperly "pass" or "fail" a hypothetical tighter cutpoint.

Fraction of excess emissions retained is the other way to measure clean screen effectiveness, as discussed above. This measure calculates the emissions in excess of the IM240 cutpoints for vehicles that fail the IM240. Excess emissions from vehicles passing the clean screen (false passes) are said to be lost, while excess emissions from vehicles failing the screen (true fails) are retained by the program. Using excess emissions to evaluate a clean screen program becomes a problem when we analyze the All Test sample of vehicles, in which vehicles are tested over different portions of the IM240. In addition, the excess emissions calculation is based on composite emissions, the cumulative grams of pollutant measured. Arizona allows vehicles to pass if their HC and CO emissions over Phase 2 of the IM240 are below a second set of cutpoints; some vehicles that pass on the basis of Phase 2 cutpoints may have composite emissions higher than the composite IM240 cutpoints. These vehicles would have excess emissions, even though they officially passed the IM240 test. We determine the size of these two sources of error by calculating the fraction of all excess emissions that come from vehicles that officially pass the IM240.

Results from the Random Sample

Table 2 presents a comparison of the results from the McClintock study of the Colorado clean screen pilot with our analysis of a similar program in Arizona. The table shows the number and distribution of tested vehicles by model year groups; the remote sensing pass rate based on two readings of less than 0.5% CO and 200ppm HC for each vehicle; the fraction of all vehicles that the screen falsely passes; and the excess emissions the screen retains, based on both start-up and final IM240 cutpoints.[6] Colorado requires IM240 testing on 1982 and newer vehicles, while Arizona requires the test for 1981 and newer vehicles; the first model year we include in this analysis is 1982.

Our analysis indicates that the vehicle distributions and remote sensing pass rates from the two states are very similar . However, the Arizona clean screen would not be as effective in identifying low emitters as the Colorado screen. The overall false pass rates are quite a bit higher in Arizona; 3% and 6%, depending on IM240 cutpoint, as opposed to 0.1% and 4% in Colorado. And the Arizona clean screen retains a smaller portion of the excess emissions (89% and 86% of the start-up HC and CO, respectively, as opposed to 100% and 97% in Colorado). In general, we see that although the false pass rate tends to decrease with more recent model years, the fraction of excess emissions retained also tends to decrease. Note that only 18% of the excess

6. In each state the cutpoints are lower, i.e. more strict, for newer model year vehicles; however, each state uses different cutpoints, and the model years for which the cutpoints are applicable vary. For example, in 1997 Colorado failed all model year 1986 to 1995 passenger vehicles with HC in excess of 4.0 grams per mile; Arizona failed all model year 1991 to 1995 vehicles with HC in excess of 1.2 grams per mile. The Arizona cutpoints are nearly identical to the start-up cutpoints recommended by EPA, and were not changed over the 18-month study period.

CO emissions are retained for the model year 1990 and newer vehicles in Arizona; this is due to two vehicles with very high excess CO emissions (124 and 243 gpm excess) falsely passing the clean screen.

Table 2. Comparison of results from Colorado and Arizona Clean Screens (1)
(at least 2 RSD within 365 days, CO < 0.5% and HC < 200ppm)

IM240 Cutpoints	Model Years	Number of Vehicles COL	Number of Vehicles AZ	Distribution of Vehicles COL	Distribution of Vehicles AZ	RSD Pass Rate COL	RSD Pass Rate AZ	False Pass Rate COL	False Pass Rate AZ	Excess Emissions Retained HC COL	Excess Emissions Retained HC AZ	Excess Emissions Retained CO COL	Excess Emissions Retained CO AZ
Start-up	1982-85	76	302	9%	11%	22%	22%	0.0%	6.6%	100%	93%	100%	92%
Start-up	1986-89	233	747	27%	26%	30%	34%	0.0%	3.3%	100%	91%	100%	97%
Start-up	1990+	559	1,825	64%	64%	59%	51%	0.2%	1.4%	100%	55%	86%	18%
	Total	868	2,914	100%	100%	48%	43%	0.1%	2.5%	100%	89%	97%	86%
Final	1982-85	76	302	9%	11%	22%	22%	9%	13%	94%	89%	97%	91%
Final	1986-89	233	747	27%	26%	30%	34%	6%	11%	91%	87%	95%	93%
Final	1990+	559	1,825	64%	64%	59%	51%	3%	3%	69%	66%	80%	36%
	Total	868	2,914	100%	100%	48%	43%	4%	6%	91%	87%	94%	87%

(1) All of the Colorado data come from an analysis of vehicles with valid CO and HC remote sensing measurements. The McClintock report provides similar data for a smaller sample of vehicles (594) with valid HC, CO and NOx remote sensing measurements. The clean screen retained 98% and 93% of the excess HC and CO emissions based on start-up IM240 cutpoints, and 91% and 93% of the excess emissions based on final IM240 cutpoints.

Results from the All Test Sample

Tables 3 and 4 present the results from the two samples of Arizona IM240 tests that we analyzed. The tables are similar to Table 2; however, we have added some additional data that help in comparing the different samples from the Arizona data (which were not published in the McClintock report on the Colorado clean screen). First, we show more detail in terms of model year groups. Second, we include the false pass rate as a percentage of all vehicles tested (as in the previous tables) and as a percentage of all vehicles that failed the IM240. This second false pass rate tells us what fraction of vehicles that failed the IM240 were falsely passed by the clean screen. Third, in addition to excess HC and CO, we calculate the excess NOx emissions retained by the clean screen, even though the Arizona clean screen does not use remote sensing NOx measurements. Finally, we show the distribution of excess emissions by model year group.

Table 3. Results from the Arizona Random Sample
(at least 2 RSD within 365 days, CO < 0.5% and HC < 200ppm)

IM240 Cut-points	Model Years	Number of Vehicles	Disn of Vehicles	Pass Rates		False Pass Rate		Excess Emissions Retained			Distribution of Excess Emissions		
				IM240	RSD	All Vehicles	Failing Vehicles	HC	CO	NOx	HC	CO	NOx
Start-up	1981	40	1%	35%	12%	5%	8%	98%	100%	77%	11%	2%	9%
Start-up	1982-85	302	10%	57%	22%	7%	16%	93%	92%	77%	49%	56%	29%
Start-up	1986-89	747	26%	85%	34%	3%	23%	91%	97%	76%	28%	32%	37%
Start-up	1990-92	677	23%	93%	42%	2%	33%	91%	82%	64%	6%	2%	17%
Start-up	1993+	1,148	39%	98%	55%	1%	36%	19%	7%	73%	6%	9%	8%
	Total	2,914	100%	88%	43%	2%	21%	89%	86%	74%	100%	100%	100 %
Final	1981	40	1%	17%	12%	10%	12%	97%	100%	85%	10%	3%	8%
Final	1982-85	302	10%	28%	22%	13%	18%	89%	91%	80%	45%	55%	29%
Final	1986-89	747	26%	59%	34%	11%	26%	87%	93%	72%	35%	33%	39%
Final	1990-92	677	23%	81%	42%	6%	31%	82%	80%	65%	7%	3%	16%
Final	1993+	1,148	39%	95%	55%	2%	36%	39%	10%	71%	4%	6%	7%
	Total	2,914	100%	75%	43%	6%	25%	87%	87%	74%	100%	100%	100 %

We make several important observations from Table 3. First, nearly 40% of the vehicles tested are model year 1993 or newer; these vehicles account for between 6% (HC) and 9% (CO) of excess emissions using the start-up cutpoints. Second, although both IM240 and remote sensing pass rates are higher for newer vehicles, the overall IM240 pass rate (88%) is much higher than that for remote sensing (43%), indicating that the remote sensing cutpoints are more stringent than those of the IM240. Relatively strict remote sensing cutpoints are desirable, since one wants to use clean screen criteria that select only the cleanest vehicles in the fleet. Third, although the false pass rates based on <u>all</u> vehicles are lower for newer model years, the false pass rates based on <u>IM240-failing</u> vehicles only are higher. This means that the clean screen passes more of the newer vehicles, but is less accurate in predicting their IM240 results. This is confirmed by examining the fraction of excess emissions retained; the clean screen retains less than 20% of the HC and CO excess emissions for the newest vehicles. The final columns indicate that these vehicles account for about 10% of total excess emissions. Finally, the Arizona clean screen retains about 70% of the excess NOx emissions, even though remote sensing measurements of NOx are not used in the screen.

As discussed above, the number of vehicles in the Random sample is relatively small, raising concerns about the representativeness of the sample. The small fraction of excess CO and HC emissions retained from model year 1993 and newer vehicles are the result of two extremely high emitters that falsely passed the screen.[7]

Table 4 presents similar data from the analysis of all IM240 tests. As one can see, this sample is much larger than the Random sample. In general, the data in Tables 3 and 4 are quite similar, suggesting that the vehicles in the Full Test sample are representative of the Arizona I/M fleet, and that the emissions adjustments used to predict full test emissions for the vehicles fast-passed or fast-failed in the All Test sample are reasonable. Note that much more of the excess

7. The excess HC/CO/NOx emissions of these two vehicles are 8.7/243/0 and 4.0/124/0 gpm.

emissions from 1993 and newer vehicles are retained in the All Test sample (69% and 65% for HC and CO, respectively) than in the Random sample (19% and 7%).

Table 4. Results from the Arizona All Test Sample
(at least 2 RSD within 365 days, CO < 0.5% and HC < 200ppm)

IM240 Cut-points	Model Years	Number of Vehicles	Disn. of Vehicles	Pass Rates IM240	Pass Rates RSD	False Pass Rate All Vehicles	False Pass Rate Failing Vehicles	Excess Emissions Retained HC	Excess Emissions Retained CO	Excess Emissions Retained NOx	Distribution of Excess Emissions HC	Distribution of Excess Emissions CO	Distribution of Excess Emissions NOx
Start-up	1981	1,795	1%	63%	17%	4%	10%	97%	98%	82%	7%	5%	4%
Start-up	1982-85	18,204	11%	67%	24%	5%	15%	92%	94%	75%	47%	53%	36%
Start-up	1986-89	42,647	25%	87%	34%	3%	19%	88%	89%	76%	35%	33%	37%
Start-up	1990-92	39,024	23%	93%	44%	2%	28%	81%	80%	73%	9%	7%	17%
Start-up	1993+	66,403	40%	99%	55%	0%	33%	69%	65%	67%	2%	2%	6%
	Total	168,073	100%	91%	43%	2%	20%	90%	91%	75%	100%	100%	100%
Final	1981	1,795	1%	38%	17%	8%	13%	95%	96%	82%	6%	5%	3%
Final	1982-85	18,204	11%	40%	24%	12%	19%	89%	92%	76%	43%	49%	31%
Final	1986-89	42,647	25%	59%	34%	12%	28%	83%	86%	73%	40%	36%	41%
Final	1990-92	39,024	23%	74%	44%	10%	39%	74%	75%	68%	10%	8%	18%
Final	1993+	66,403	40%	91%	55%	5%	49%	62%	59%	59%	2%	3%	7%
	Total	168,073	100%	73%	43%	8%	31%	85%	88%	72%	100%	100%	100%

As discussed above, the All Test sample is valuable since it includes many more vehicles than the Random sample, and likely is most representative of the on-road fleet. Because vehicles are tested over different portions of the IM240 cycle, however, the emissions results of individual vehicles are not directly comparable. How accurate are our simple adjustments to predict full IM240-equivalent emissions for the vehicles that are fast passed or fast failed? Table 5 shows what portion of excess adjusted emissions are attributable to vehicles that officially pass the IM240, from the All Test sample. Here we see that our adjustments to the fast pass/fast fail IM240 test results incorrectly assign as much as 3% of the overall excess emissions to vehicles that actually passed the IM240. That is, our adjustments overestimate full IM240 emissions for a relatively small portion of passing vehicles; the overestimation is greater for CO than for HC or NOx. However, our adjustments appear less accurate for the newest vehicles; 25% of the excess CO emissions from 1993 and newer vehicles are overestimated full IM240 emissions from fast-passed vehicles. [8]

8. There are no excess emissions from passing vehicles using the final cutpoints, because the cutpoints themselves are used to determine whether an individual vehicle passes or fails the IM240. Similarly, there are no excess emissions from passing vehicles from the Random sample, because no adjustments are needed to predict full IM240 emissions (although it is possible to have excess emissions from vehicles that fail composite cutpoints but pass Phase 2 cutpoints, since the excess emissions are calculated based on the composite IM240 emissions).

Table 5. Fraction of "Excess Emissions" from Vehicles Passing the IM240

IM240 Cutpoints	Model Year	"Excess Emissions" from Passing Vehicles HC	CO	NOx
Start-up	1981	0%	0%	0%
Start-up	1982-85	1%	1%	0%
Start-up	1986-89	2%	3%	0%
Start-up	1990-92	3%	8%	1%
Start-up	1993+	5%	25%	1%
	Total	1%	3%	0%

The large number of vehicles in the All Test sample allows us to examine the relative accuracy of the clean screen in identifying low-emitting cars versus light duty trucks. Table 6 shows the statistics for cars and light duty trucks of all model years (heavier trucks, with gross vehicle weights greater than 6,000 pounds and subject to looser emissions cutpoints than lighter trucks, are included in our truck category). A larger fraction of trucks pass the IM240 than cars (93% versus 89%); this suggests that the IM240 cutpoints for trucks are less stringent than those for cars. Interestingly, the same remote sensing cutpoints applied to all vehicle types result in identical pass rates for cars and trucks (43%). The false pass rates and percent excess emissions retained are nearly identical for cars and trucks, with the exception of the amount of excess emissions retained based on final cutpoints.

Table 6. Results from the Arizona All Test Sample
(at least 2 RSD within 365 days, CO < 0.5% and HC < 200ppm)

IM240 Cutpoints	Type	Number of Vehicles	Pass Rates IM240	RSD	False Pass Rate All Vehicles	Failing Vehicles	Excess Emissions Retained HC	CO	NOx
Start-up	Cars	102,764	89%	43%	2%	20%	89%	90%	75%
Start-up	LDTs	65,309	93%	43%	1%	19%	90%	94%	75%
Final	Cars	102,764	69%	43%	10%	31%	83%	76%	82%
Final	LDTs	65,309	78%	43%	7%	30%	87%	91%	71%

Table 7 presents the summary results for all model year 1982 and newer vehicles, from the Colorado study and the two Arizona samples (the data from the Colorado study and the Arizona Random sample are taken from Table 2; the Arizona All Test data are taken from Table 4, excluding model year 1981 vehicles). The table indicates that the Colorado pilot clean screen appears to be more effective in identifying clean vehicles than a similar program applied to the Random and All Test samples of vehicles in Arizona. Differences in the effectiveness of identifying clean vehicles in the two states may be due to a variety of factors:

- the small sample size of the Colorado study, relative to the two Arizona data samples;
- differences in remote sensing measurement and/or IM240 measurement technology and techniques; and
- regional differences (how different weather, fuels, maintenance practices, and driving habits affect in-use emissions).

Table 7. Summary Results for MY82+ Vehicles
(at least 2 RSD within 365 days, CO < 0.5% and HC < 200ppm)

Sample	IM240 Cutpoints	Vehicles	RSD Pass Rate	False Pass Rate	Excess Emissions Retained HC	Excess Emissions Retained CO
Colorado	Start-up	868	48%	0.1%	100%	97%
Arizona Random	Start-up	1,726	43%	2.5%	89%	86%
Arizona All Test	Start-up	99,875	36%	2.8%	89%	91%
Colorado	Final	868	48%	4%	91%	94%
Arizona Random	Final	1,726	43%	6%	87%	87%
Arizona All Test	Final	99,875	36%	11%	85%	88%

Results of Model Year Exemptions

In the previous analysis we have seen that clean screening is less effective in identifying the cleanest vehicles among newer vehicles. Rather than applying a clean screen, some states may choose to exempt from IM240 testing all vehicles of the newest model years.[9] Because Arizona currently does not exempt the newest vehicles from testing, we examined how effective model year exemptions are in terms of retaining excess IM240 emissions.[10]

Table 8 presents the model year exemption analysis for the Arizona Random and All Test samples, based on the start-up IM240 cutpoints. For each sample, we calculate the cumulative fraction of vehicles and excess emissions by model year. Reading down the columns, one can determine the fraction of vehicles exempted for any group of model years, and the excess emissions associated with those vehicles. For example, 17% of all vehicles in the Random sample are model years 1995 and newer; Arizona could exempt all these vehicles from I/M testing and still not lose any of the excess emissions identified by the IM240 program.

The last row of Table 8 shows the results of applying the clean screen to all 1981 and newer vehicles in both samples, taken from Tables 3 and 4 (Table 8 shows excess emissions lost, which is the inverse of excess emissions retained in the earlier tables). In all cases, the same portion of the fleet could be exempted based on vehicle model year, and fewer excess emissions would be lost. For example, clean screening the All Test sample would relieve 43% of the vehicles in the Arizona fleet from testing, while losing 9% to 25% of the excess IM240 emissions, depending on the pollutant. On the other hand, exempting model year 1992 and newer vehicles would relieve 47% of vehicles from testing, with losses in excess emissions of only 4% to 11% (using the All Test sample). Model year exemptions are even more effective when one considers that all registered vehicles of the exempted model years would automatically be relieved of testing. On the other hand, only those vehicles that drive by a remote sensor and pass the emissions cutpoints would be exempt from testing under a clean screen (recall that less than one-half of the vehicles in the Arizona fleet were matched with valid remote sensing readings, as discussed above).

9. For example, in late 1994 Colorado began exempting vehicles from the four newest model years from I/M testing.
10. Arizona does allow owners of new vehicles to opt out of the first I/M testing cycle. However, the state does not have an estimate for the portion of eligible vehicles that do so. Arizona began exempting the newest five model years from their first I/M test in late 1998.

Table 8. Cumulative Vehicle Distibutions and Excess Emissions, by Model Year and Arizona IM240 Sample

	Random Sample				All Test Sample			
	Vehicle	Cumulative Excess Emissions			Vehicle	Cumulative Excess Emissions		
Model Year	Disn.	HC	CO	NOx	Disn.	HC	CO	NOx
1997	0%	0%	0%	0%	0%	0%	0%	0%
1996	4%	0%	0%	0%	4%	0%	0%	0%
1995	17%	0%	0%	0%	17%	0%	0%	1%
1994	29%	2%	3%	1%	29%	1%	1%	2%
1993	39%	6%	9%	8%	40%	2%	2%	6%
1992	*47%*	*9%*	*9%*	*12%*	*48%*	*5%*	*4%*	*11%*
1991	55%	11%	10%	19%	56%	8%	7%	18%
1990	63%	12%	11%	25%	63%	11%	9%	23%
1989	70%	16%	15%	34%	70%	15%	13%	31%
1988	76%	21%	20%	41%	77%	23%	20%	43%
1987	82%	28%	30%	48%	82%	33%	31%	51%
1986	88%	40%	42%	62%	88%	46%	43%	61%
1985	92%	54%	57%	74%	92%	63%	61%	73%
1984	96%	71%	79%	82%	96%	78%	78%	84%
1983	97%	81%	94%	90%	98%	87%	90%	91%
1982	99%	89%	98%	91%	99%	93%	95%	96%
1981	100%	100%	100%	100%	100%	100%	100%	100%
Clean Screen (1981+)	*43%*	*11%*	*14%*	*26%*	*43%*	*10%*	*9%*	*25%*

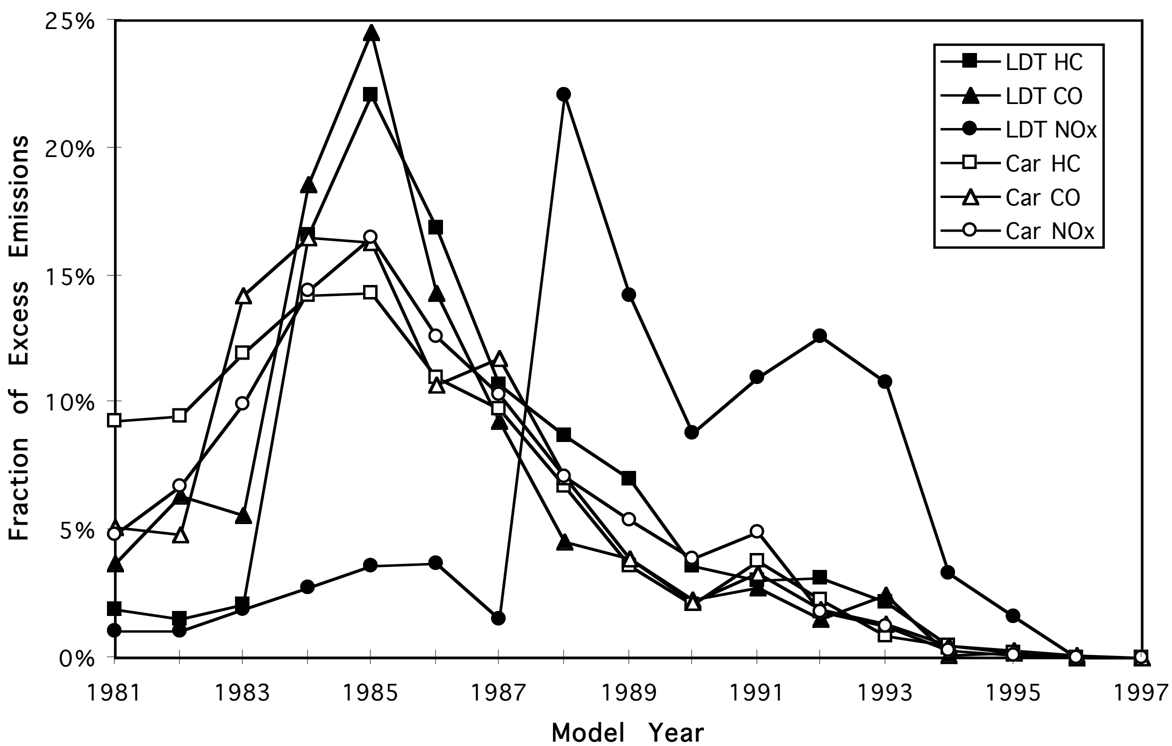
Table 9 presents the effect of model year exemptions for passenger cars and light duty trucks, based on the Arizona All Test sample. Again, exempting model year 1992 and newer cars from I/M testing results in a smaller loss of excess emissions (3% or 4%) than using a remote sensing clean screen (10% to 25%). However, exempting 1992 and newer trucks from I/M testing results in a slightly larger loss of excess NOx emissions (28%) than the clean screen (25%), even though the Arizona clean screen is based on HC and CO measurements only.

Figure 1 shows the distribution of excess emissions (based on start-up IM240 cutpoints) by model year and vehicle type, using the All Test sample (the emissions distributions from the Random sample are similar). The distribution of excess emissions for the three pollutants is quite similar for cars and trucks, with the exception of truck NOx emissions. A large fraction of excess NOx emissions from trucks comes from model year 1988 and newer trucks. The difference in the car and truck excess NOx distributions may be due to the relative stringency of the standards for each: 1988 and newer trucks are subject to much tighter NOx cutpoints than earlier trucks (3.5 gpm NOx for 1988 and newer trucks under 6,000 pounds gross vehicle weight, as opposed to 7.0 gpm NOx for 1987 and older trucks).

Table 9. Cumulative Vehicle Distibutions and Excess Emissions, by Model Year and Vehicle Type

	Cars				Light Duty Trucks			
	Vehicle	Cumulative Excess Emissions			Vehicle	Cumulative Excess Emissions		
Model Year	Disn.	HC	CO	NOx	Disn.	HC	CO	NOx
1997	0%	0%	0%	0%	0%	0%	0%	0%
1996	4%	0%	0%	0%	4%	0%	0%	0%
1995	15%	0%	0%	0%	19%	0%	0%	2%
1994	26%	1%	1%	0%	34%	1%	0%	5%
1993	36%	2%	2%	2%	45%	3%	3%	16%
1992	*44%*	*4%*	*4%*	*3%*	*54%*	*6%*	*4%*	*28%*
1991	52%	8%	7%	8%	61%	9%	7%	39%
1990	60%	10%	10%	12%	67%	13%	9%	48%
1989	68%	13%	13%	18%	74%	20%	13%	62%
1988	75%	20%	21%	25%	80%	28%	18%	84%
1987	81%	30%	32%	35%	85%	39%	27%	86%
1986	87%	41%	43%	48%	90%	56%	41%	90%
1985	91%	55%	59%	64%	94%	78%	66%	93%
1984	95%	69%	76%	79%	97%	95%	84%	96%
1983	97%	81%	90%	89%	98%	97%	90%	98%
1982	99%	91%	95%	95%	99%	98%	96%	99%
1981	100%	100%	100%	100%	100%	100%	100%	100%
Clean Screen (1981+)	*43%*	*11%*	*10%*	*25%*	*43%*	*10%*	*6%*	*25%*

Figure 1. Distribution of Excess Start-Up Emissions by Model Year and Vehicle Type, All Test Sample, 1996-97 Arizona IM240



With the exception of truck NOx emissions, exempting the newest model years of vehicles from I/M testing appears to be more effective than applying a clean screen based on remote sensing measurements. However, there are reasons why a state may not want to exempt the newest model years from I/M testing. One reason is the sense of equity; motorists may feel that all vehicles should be given the same opportunity to fail an I/M test. Another reason is that a state may wish to test a large representative sample of vehicles do demonstrate to EPA the effectiveness of their I/M program. Finally, a state may want to ensure that vehicles get at least one I/M inspection before the vehicle's warranties on emissions-related components expire. States should consider the number of vehicles excused from I/M testing, as well as the expected loss in emission reductions, when deciding between using a remote sensing based clean screen or model year exemptions.

Other Criteria for Remote Sensing Data

As discussed above, for this study we applied the Colorado clean screen cutpoints to the last two remote sensing measurements of individual vehicles in Arizona. In order to test the sensitivity of our results to how we used the remote sensing data for individual vehicles, we used the remote sensing data in three other ways:

1) the last single reading prior to the IM240 test ("Last RSD");
2) the average of all multiple readings prior to the IM240 test ("Avg of All"); and
3) the average of the last two readings prior to the IM240 test ("Avg of Last 2").

Table 10 shows the absolute difference in the results from those obtained for all vehicles, as shown in Tables 3 and 4. For example, the clean screen based on the last remote sensing measurement ("Last RSD") using the random IM240 sample ("Random") resulted in a higher percentage of vehicles passing the screen and a higher false pass rate, and less excess emissions retained (5, 14 and 11 percentage points less HC, CO, and NOx, respectively). Table 10 indicates that the three other ways of using remote sensing data are less effective than using both of the last two remote sensing measurements: more vehicles pass the remote sensing screen under the other scenarios, yet false pass rates are higher and a smaller percentage of the excess emissions are retained. On the other hand, using the last measurement ("Last RSD") and the average of all measurements ("Avg of All") requires only one remote sensing reading per vehicle, while the other two scenarios require at least two measurements. As demonstrated above, requiring at least two readings substantially reduces the remote sensing coverage of the fleet.

Table 10. Absolute Difference from Comparable Last 2 RSD Scenario

IM240 Sample	RSD Scenario	IM240 Cutpoints	RSD Pass Rate	False Pass Rate	Excess Emissions Retained		
					HC	CO	NOx
Random	Last RSD	Start-up	+19%	+10%	-5%	-4%	-11%
Random	Avg of All	Start-up	+18%	+14%	-11%	-11%	-14%
Random	Avg of Last 2	Start-up	+17%	+10%	-6%	-5%	-12%
All	Last RSD	Start-up	+18%	+11%	-8%	-7%	-14%
All	Avg of All	Start-up	+18%	+15%	-11%	-11%	-18%
All	Avg of Last 2	Start-up	+18%	+11%	-7%	-6%	-14%
Random	Last RSD	Final	+19%	+5%	-8%	-6%	-13%
Random	Avg of All	Final	+18%	+19%	-13%	-13%	-16%
Random	Avg of Last 2	Final	+17%	+14%	-8%	-8%	-13%
All	Last RSD	Final	+18%	+15%	-10%	-9%	-15%
All	Avg of All	Final	+18%	+18%	-14%	-12%	-18%
All	Avg of Last 2	Final	+18%	+15%	-9%	-7%	-14%

As discussed above, we used remote sensing measurements made up to one year prior to the vehicle's initial I/M test. Since vehicle owners may make changes to their vehicles prior to bringing them in for I/M testing, allowing such a long period between remote sensing and I/M measurement may account for some of the inaccuracy in the clean screen. We did not test the sensitivity of the clean screen to shortening the time allowed between remote sensing measurement and I/M test. However, Table 11 shows how the clean screen coverage would be affected by reducing the time between the two measurements. As the table indicates, only half of the vehicles with valid remote sensing and IM240 measurements in the Random sample were measured by the remote sensor within 90 days prior to I/M testing (the fraction is slightly smaller if two remote sensing readings are required). The results from Table 11 can be combined with the calculated fleet coverage based on up to one year between remote sensing and I/M testing (from Table 1) to determine the impact on overall fleet coverage. If only remote sensing measurements within 90 days of I/M testing are used, the overall clean screen coverage would be reduced by one-half, to only 15% of the I/M fleet (or 8%, if at least two remote sensing readings are required). The dramatic reduction in coverage from limiting the time between remote sensing and I/M measurement will dampen any improvement in the accuracy of the clean screen.

Table 11. Fraction of Vehicles Measured by Remote Sensing, by Time between Measurement and I/M Test, Arizona Random Sample

	At Least 1 RSD Reading		At Least 2 RSD Readings	
Number of Days between RSD and I/M Measurement	Cumulative Number of Vehicles	Cumulative Percent of Vehicles	Cumulative Number of Vehicles	Cumulative Percent of Vehicles
30	1,064	23%	505	17%
60	1,764	38%	901	31%
90	2,311	50%	1,246	43%
120	2,762	59%	1,560	53%
180	3,476	75%	2,042	70%
365	4,651	100%	2,916	100%

Conclusions

In this report we have applied the methodology and cutpoints of a pilot clean screen program in Colorado to the Arizona I/M area, using 18 months of remote sensing and IM240 data. Our primary conclusions are:

- Clean screening can be an effective method to exempt a fraction of the vehicle fleet from I/M testing, while retaining nearly all of the excess emissions. A clean screen program can make an I/M program more cost-effective by concentrating resources on vehicles that are more likely to be high emitters. By exempting the cleanest vehicles from regular I/M testing, a clean screen program can also reduce the inconvenience to the public of bringing likely clean vehicles in for I/M testing.

- A small fraction of the I/M fleet was measured by remote sensing in Arizona; about one-third of the vehicles were matched with at least one remote sensing measurement, and only 20% had at least two measurements. Such low coverage reduces the effectiveness of using remote sensing to exempt vehicles from I/M testing. On the other hand, there was an incentive for Arizona drivers to avoid driving by the remote sensors. A true clean screen program, where there is no potential penalty for being measured by a remote sensor, could encourage more drivers to drive by the sensors, resulting in greater fleet coverage.

- The clean screen program applied to Arizona vehicles is not as effective as the same program applied to Colorado vehicles; for example, using start-up IM240 cutpoints, the Colorado clean screen retains 97% of excess HC, and 100% of excess CO emissions, while the clean screen applied to the Arizona Random sample retains only 93% and 94% of excess emissions (Table 8). This result holds even after we exclude the newest vehicles that are exempted from IM240 testing in Colorado, and examine a larger sample of vehicles from Arizona that is more representative of the I/M fleet than the Random sample of full IM240 tests. It is likely that the noted problems with the Arizona remote sensing measurements affect the accuracy of the Arizona clean screen. Additional differences in how the two states measure emissions (using both remote sensing and IM240), the small samples of vehicles studied in the Colorado study and the Arizona Random sample, and regional differences that may affect in-use emissions (such as weather, fuels, maintenance practices, and driving habits) also likely account for the differences in the effectiveness of clean screen programs in the two states.

- Clean screening using remote sensing is not as accurate for newer vehicles as it is for older vehicles. This may be due to several reasons, including the relatively few high emitters among the most recent model years, and the use of a single remote sensing cutpoint for all ages of vehicles. The accuracy of the screen in identifying clean vehicles from later model years may improve as this cohort of vehicles ages. Tighter remote sensing cutpoints applied to newer vehicles may also improve the clean screen accuracy for newer vehicles.

- Remote sensing clean screening is just as accurate for light duty trucks as it is for passenger cars, even if the same remote sensing cutpoints are used for different vehicle types.

- Exempting entire model years of <u>cars</u> from initial I/M testing is more effective than using a remote sensing clean screen for the newest vehicles. Model year exemptions also apply to all vehicles of the model years exempted, and not just the clean vehicles that drive by a remote sensor. However, it is likely that a clean screen that uses remote sensing NOx measurements would be more effective than a model year exemption of <u>light duty trucks</u>, since nearly 30% of excess NOx emissions from trucks comes from relatively recent model years (1992 and newer). In addition, there are other reasons why states may choose not to exempt new vehicles from I/M testing.

- Applying cutpoints to the last two remote sensing readings of an individual vehicle appears to be the best use of remote sensing data to predict that vehicle's IM240 result. However, this approach requires at least two remote sensing readings per vehicle, which reduces the remote sensing coverage of the vehicle fleet.

- Restricting the time between remote sensing measurement and the next scheduled I/M test to 90 days results in a reduction in fleet coverage of 50%.

In this report we have not examined the sensitivity of various aspects of the Colorado pilot on clean screen accuracy. These include: using alternative remote sensing cutpoints; shortening the time allowed between remote sensing measurement and IM240 test; or restricting the analysis to measurements made during stabilized vehicle behavior, based on average site characteristics or speed and acceleration measurements of individual vehicles. This report also does not analyze the effectiveness of using remote sensing to identify suspected high emitters. We plan to do these types of analysis in the future.

References

1. McClintock, Peter M. 1998. *The Colorado Enhanced I/M Program 0.5% Sample Annual Report.* Prepared for the Colorado Department of Public Health and Environment. January 27.

2. Glover, Ed. 1996. *Analysis of Arizona Remote Sensing Data*. Draft memo to Phil Lorang, US EPA. August.

3. Radian International LLC. 1996. *Exemption of Vehicles from I/M Requirements in the Phoenix Area: The Arizona Clean Screen Program*. Prepared for the Arizona Department of Environmental Quality. December 4.

4. Wenzel, Tom, Marc Ross, and Robert Sawyer. 1997. *Analysis of Emissions Deterioration of In-Use Vehicles: Presentation at the 1997 SAE Government/Industry Meeting*. Lawrence Berkeley National Laboratory, LBL-41452. May.

5. Heirigs, Philip L. and Jay Gordon. 1996. *Preconditioning Effects on I/M Test Results Using IM240 and ASM Procedures*. SAE Technical Paper Series 962091. October.

6. Rendahl, Craig S. 1997. "Intercomparisons between Smog Dog RES and RSTi RSD," internal report. November 11.

7. Klausmeier, Rob, Sandeep Kishan, Rick Baker, Andrew Burnette, and Joel McFarland. 1995. *Draft Final Report: Evaluation of the California Pilot Inspection/Maintenance (I/M) Program*. Prepared for the California Bureau of Automotive Repair. March 31.

8. Klausmeier, Rob and Peter McClintock. 1997. *The Greeley Remote Sensing Pilot Program, Final Report.* Prepared for the Colorado Department of Public Health and Environment. November 4.

APPENDIX C

Applying Clean Screen Remote Sensing Program to Heavy Duty Gasoline Vehicles

Tom Wenzel, Lawrence Berkeley National Laboratory

March 12, 1999

I merged the July through December 1997 Arizona idle I/M test result data with remote sensing data, and got 4,156 heavy duty gasoline vehicles (HDTs) with both RSD and idle tests (or 16% of the 26,000 HDTs given idle tests). The coverage is low, because I used only 6 months of idle data (as opposed to 18 months for the matching with IM240 data). Figure 1 shows the distribution of matched vehicles by model year. Only 2,471 HDTs (or 10% of all HDTs) have more than one RSD reading.

Figures 2 through 5 show idle, loaded idle, and RSD measurements on the same vehicles from the last 6 months of 1997 in Arizona. RSD measurements are the average of all readings for a particular vehicle up to 365 days prior to initial I/M test. Figures 2 and 3 show that RSD CO and HC are higher than idle CO and HC for HDTs; RSD HC for older HDTs is much higher than idle HC. Both CO and HC RSD curves show an increase in emissions for MY94 and MY95 HDTs.

Figures 4 and 5 compare idle, loaded idle, and RSD measurements on light-duty vehicles. Here old vehicles have about the same idle and RSD emissions, but new vehicles have much higher RSD emissions than idle emissions.

Applying the same RSD cutpoints as I used in the analysis of light duty vehicles, 200 ppm HC and 0.5 % CO, results in 24% of HDTs being excused from I/M testing, while retaining 93% of excess idle HC and 94% of excess idle CO, and retaining 100% of excess loaded idle HC and 95% of excess loaded idle CO. Table 1 shows the effect of exempting whole model years of HDTs from testing. If MY90 and newer trucks are exempted from I/M testing, the program would test 60% of the trucks (i.e. excusing 40% from testing), while retaining 98% and 95% excess idle HC and CO, and 93% and 91% excess loaded HC and CO.

Figure 1. Number of Vehicles with RSD Reading, by Model Year and Type

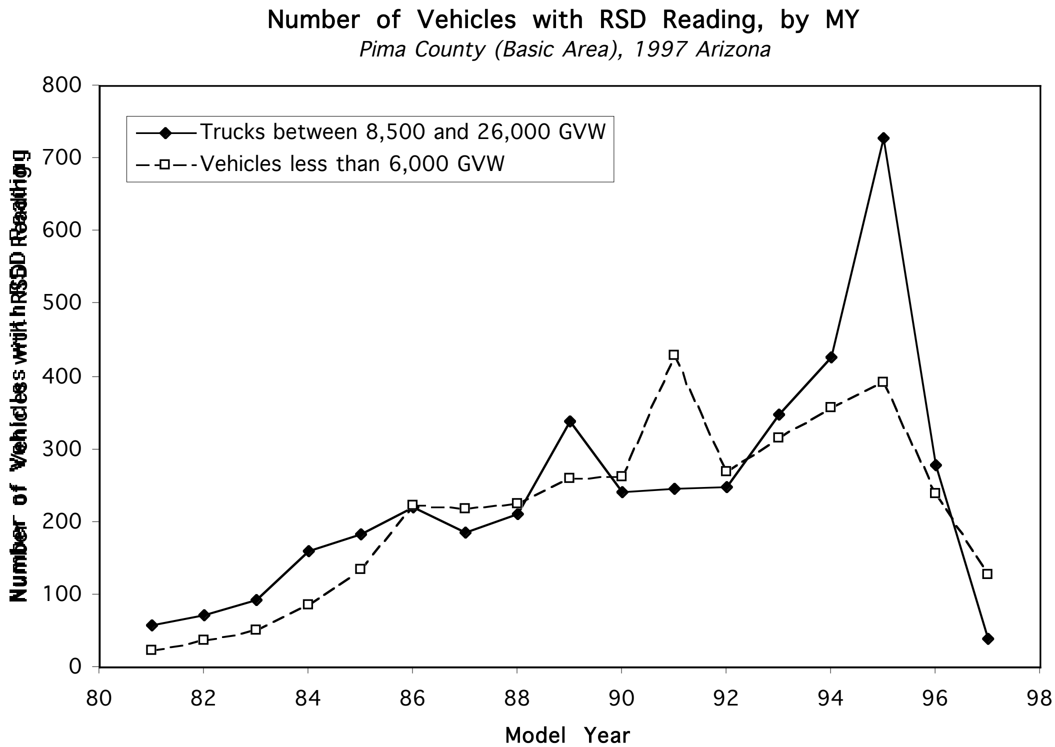


Figure 2. Average CO by Model Year, Heavy Duty Gasoline Vehicles

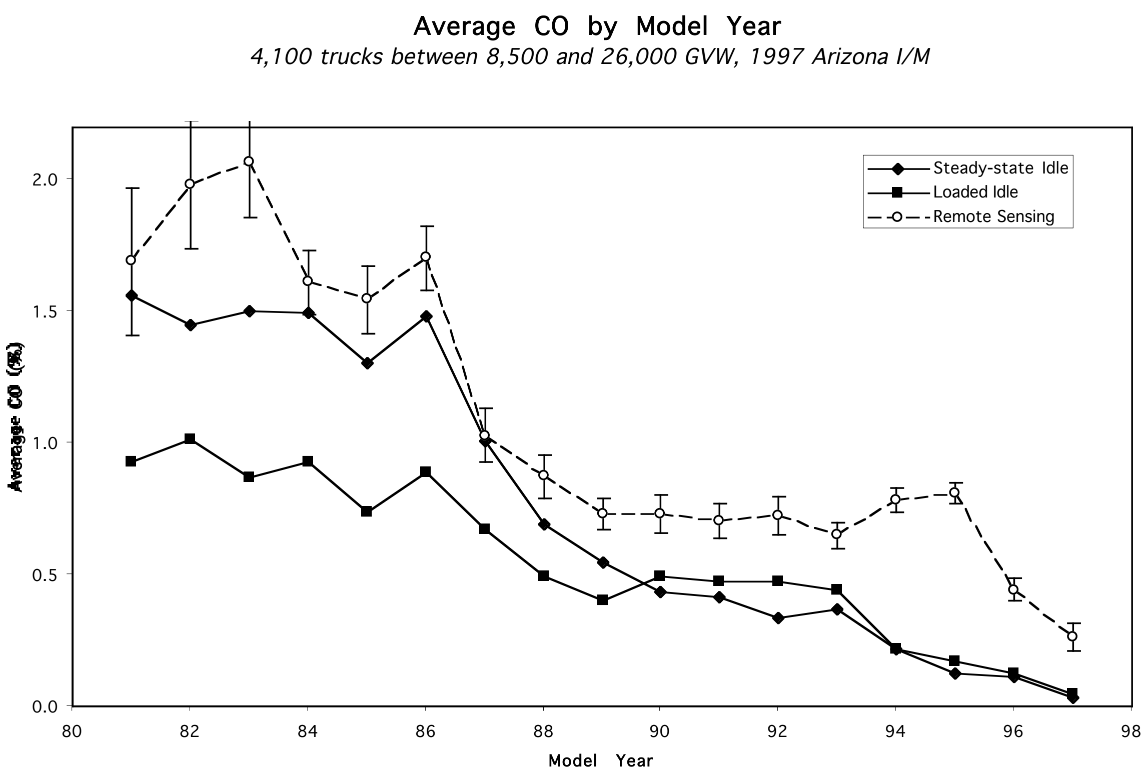


Figure 3. Average HC by Model Year, Heavy Duty Gasoline Vehicles

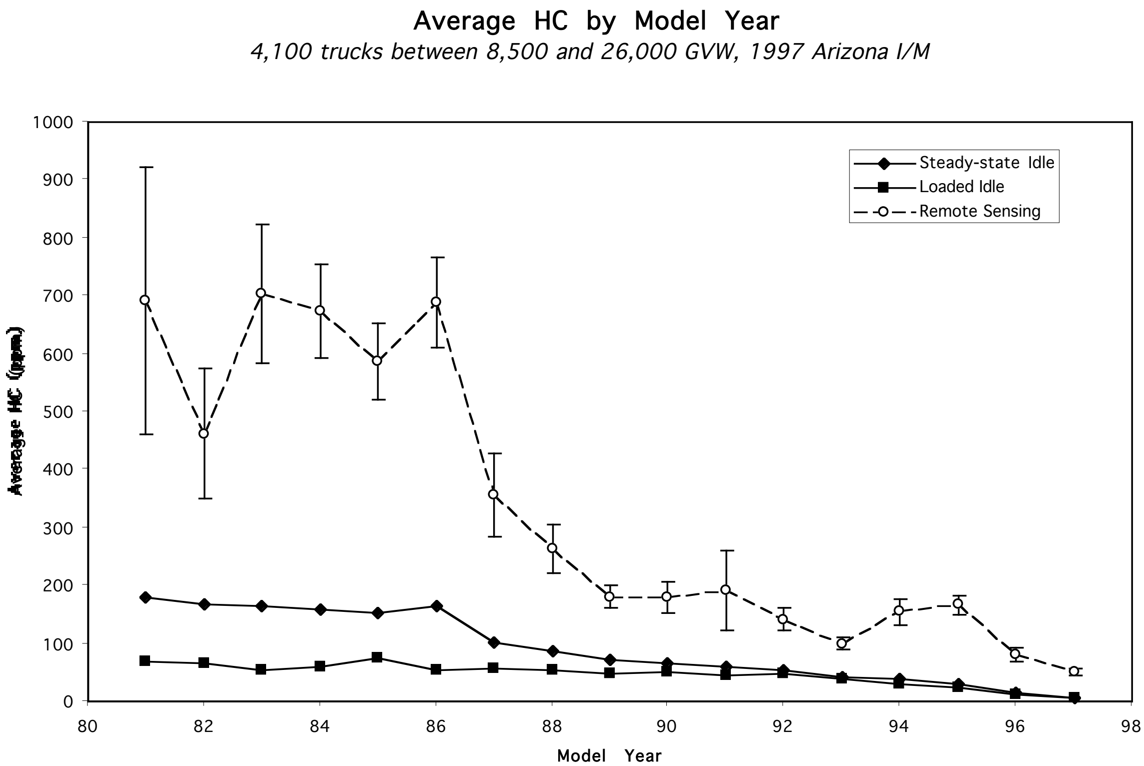


Figure 4. Average CO by Model Year, Light Duty Vehicles

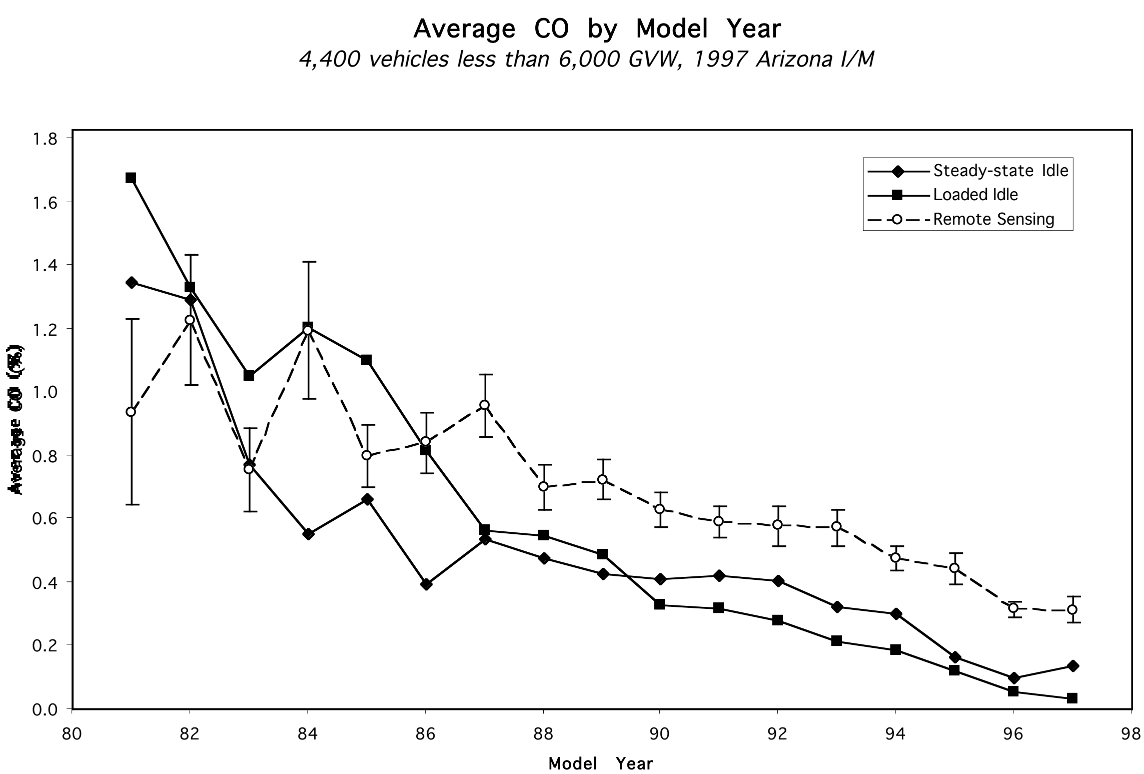


Figure 5. Average HC by Model Year, Light Duty Vehicles

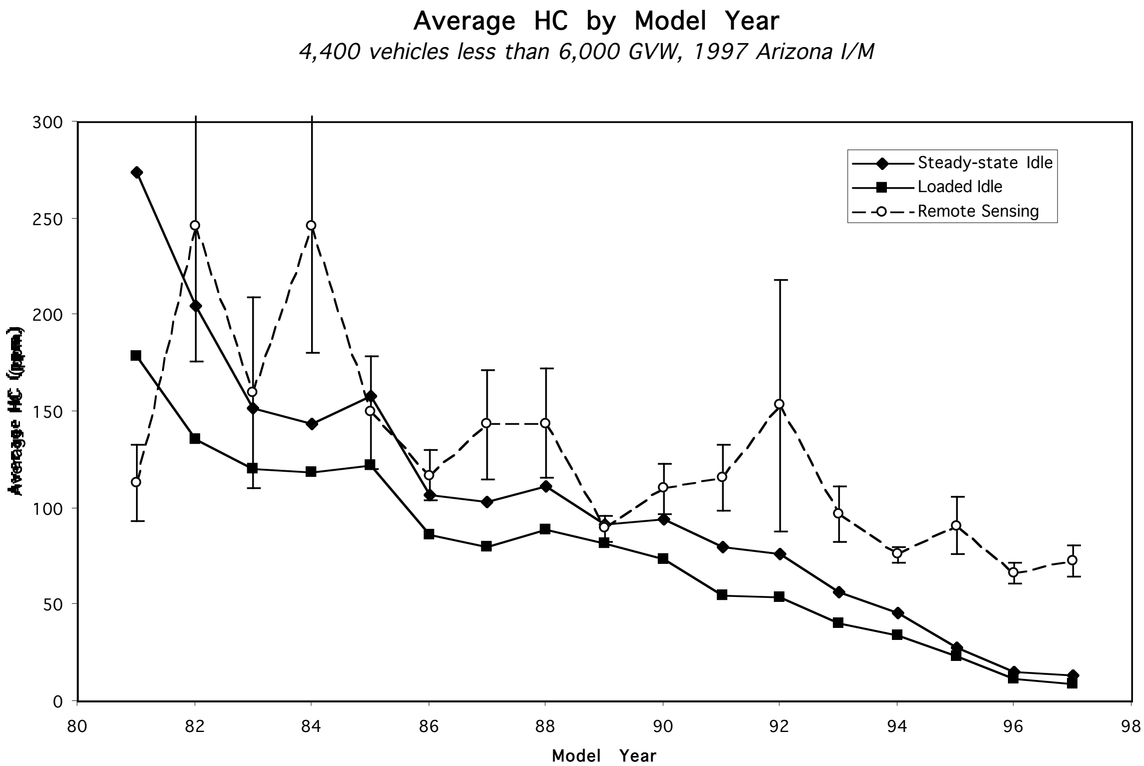


Table 1. Distribution of Initial Idle Test Emissions of 29,000 Gasoline Trucks between 8,500 and 26,000 GVW

July-December 1997 Arizona I/M

MY	Cumulative Distribution				
	Vehicles	Excess Idle		Excess Loaded	
		HC	CO	HC	CO
67	0%	0%	0%	0%	1%
68	0%	0%	0%	0%	1%
69	0%	0%	1%	0%	2%
70	1%	1%	1%	0%	3%
71	1%	2%	3%	0%	4%
72	1%	3%	4%	1%	5%
73	2%	6%	6%	2%	7%
74	3%	7%	8%	3%	9%
75	4%	8%	9%	3%	10%
76	5%	11%	14%	4%	14%
77	6%	13%	17%	4%	17%
78	8%	16%	24%	5%	22%
79	11%	20%	30%	6%	26%
80	14%	27%	37%	9%	31%
81	16%	30%	41%	10%	34%
82	18%	37%	46%	19%	40%
83	21%	43%	51%	24%	45%
84	24%	49%	57%	28%	53%
85	29%	60%	65%	42%	62%
86	36%	74%	74%	65%	71%
87	42%	86%	85%	77%	79%
88	47%	93%	90%	87%	84%
89	53%	96%	93%	91%	88%
90	***60%***	***98%***	***95%***	***93%***	***91%***
91	65%	99%	96%	96%	94%
92	70%	99%	97%	96%	95%
93	75%	99%	97%	96%	96%
94	80%	99%	99%	96%	99%
95	86%	100%	100%	98%	100%
96	95%	100%	100%	100%	100%
97	99%	100%	100%	100%	100%
98	100%	100%	100%	100%	100%

APPENDIX D

Comparison of Emissions by Model using IM240 and Remote Sensing Data

Tom Wenzel, Lawrence Berkeley National Laboratory
July 24, 1998

This memo describes our comparison of Arizona IM240 and remote sensing failure rates and average emissions by vehicle model. We examine whether remote sensing data correlates with IM240 data, so that failure rates/average emissions by model as measured by remote sensing can be used to update low- and high-emitter profiles. We also examine what effect engine size has on remote sensing readings.

The IM240 data are from 1996. Failure rates by model are based on actual pass/fail decisions made by the Arizona IM240 contractor, and include fast passes, fast fails, and phase 2 passes. Full IM240 equivalent emissions are calculated for vehicles that fast-pass or fast-fail the I/M test, in order to calculate average emissions by model. Remote sensing measurements are from January 1996 to June 1997. The RSD contractor began measuring vehicle speed and acceleration in October 1996. About half of the vehicles with acceleration measurements were decelerating when they passed the instrument; we excluded these from the analysis (since they might exhibit high "off-cycle" HC emissions). We included all vehicles measured prior to October 1996; many of these may have been decelerating when measured by the remote sensor. We also averaged all multiple readings of the same vehicle. We use cutpoints of 1 percent CO and 200 ppm HC to determine vehicles "failing" remote sensing.

The figures show the failure rates and average emissions by model year and model. We limit the analysis to models that had both IM240 and RSD measurements of at least 100 individual vehicles (individual vehicles are not necessarily measured by both methods). The first 2 figures compare IM240 and RSD results of CO and HC failure rate by model and engine displacement. The correlation between IM240 and RSD is not good, in part because many models have very low failure rates. We identify some models with high failure rates for either CO or HC. The second figures compare the IM240 and RSD average emissions. Here we see better correlation between the two test methods. We also see that RSD tends to be biased against models with smaller engines. Small engine models have high average CO based on RSD, but relatively low average CO based on IM240; the opposite is true for models with larger engines. The bias is not as apparent for HC as for CO. This bias presumably is due to the comparison of emission concentrations measured by RSD with the mass emissions measurements of the IM240. If so, the bias would be corrected by adjusting the remote sensing concentration emissions to mass emissions, by assuming a fuel economy by vehicle age and/or model, or by converting both the remote sensing concentrations and IM240 grams per mile emissions to grams per gallon of fuel burned, using the measured CO_2 emissions.

The table gives the data for the models identified in the figures.

Figure 1. CO Fail Rate by Engine Displacement, IM240 vs. RSD

Excluding all deceleration RSD readings, MY90-93 Car Models, Arizona

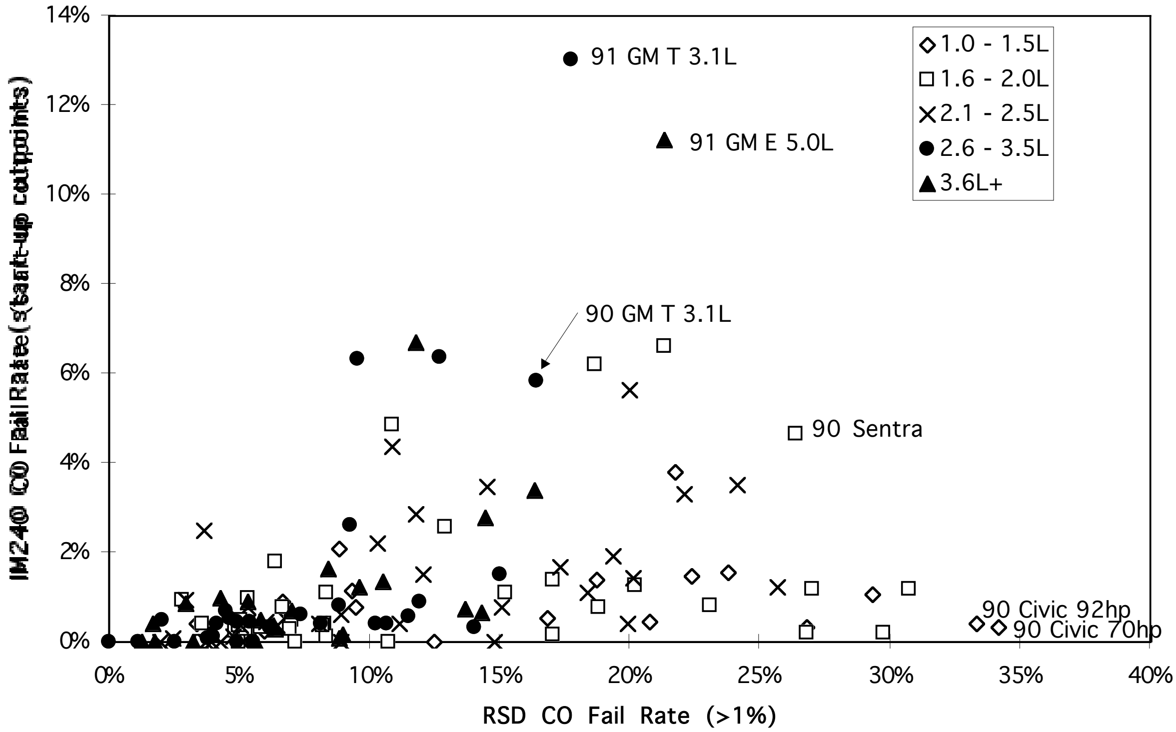


Figure 2. HC Fail Rate by Engine Displacement, IM240 vs. RSD

Excluding all deceleration RSD readings, MY90-93 Car Models, Arizona

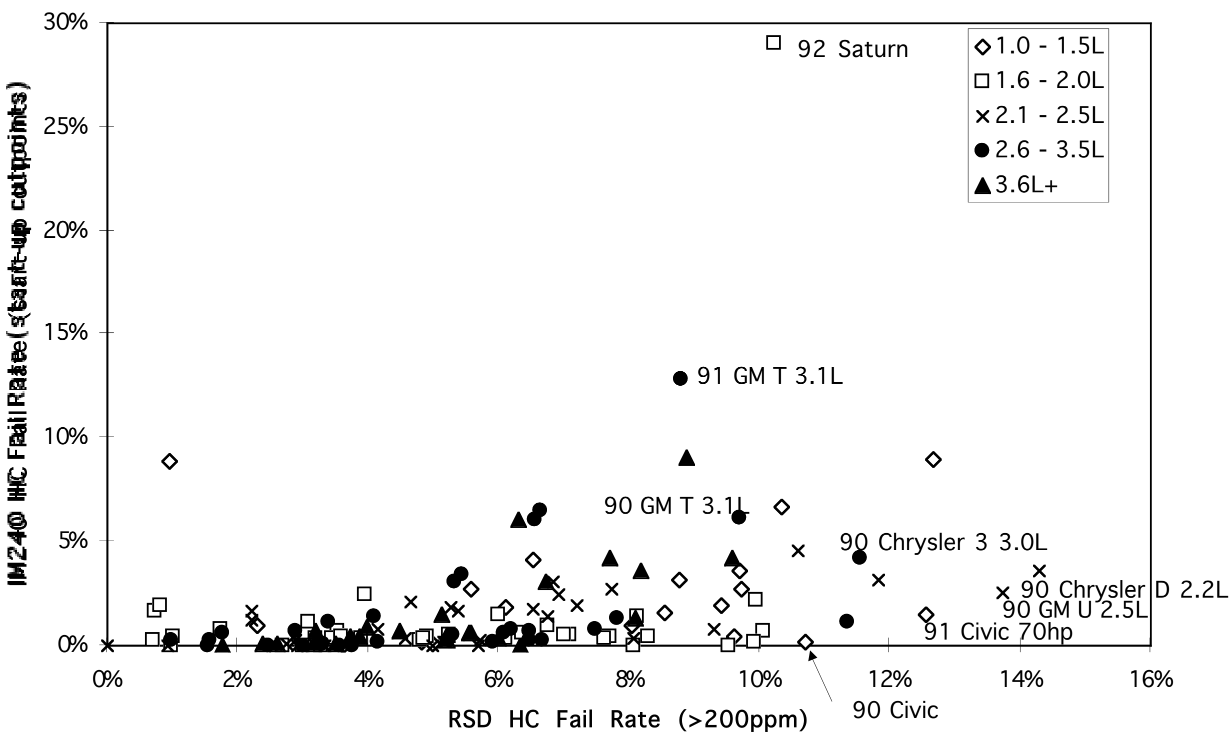


Figure 3. Average CO by Engine Displacement, IM240 vs. RSD
Excluding all deceleration RSD readings, MY90-93 Car Models, Arizona

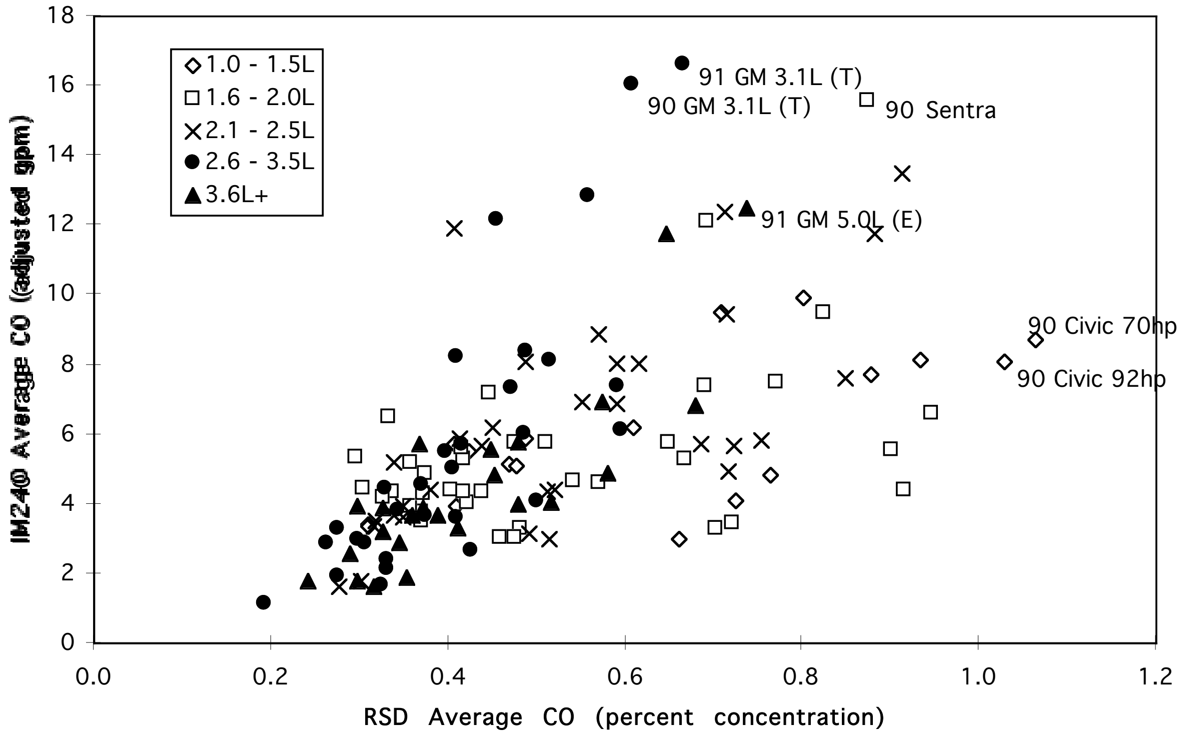


Figure 4. Average HC by Engine Displacement, IM240 vs. RSD
Excluding all deceleration RSD readings, MY90-93 Car Models, Arizona

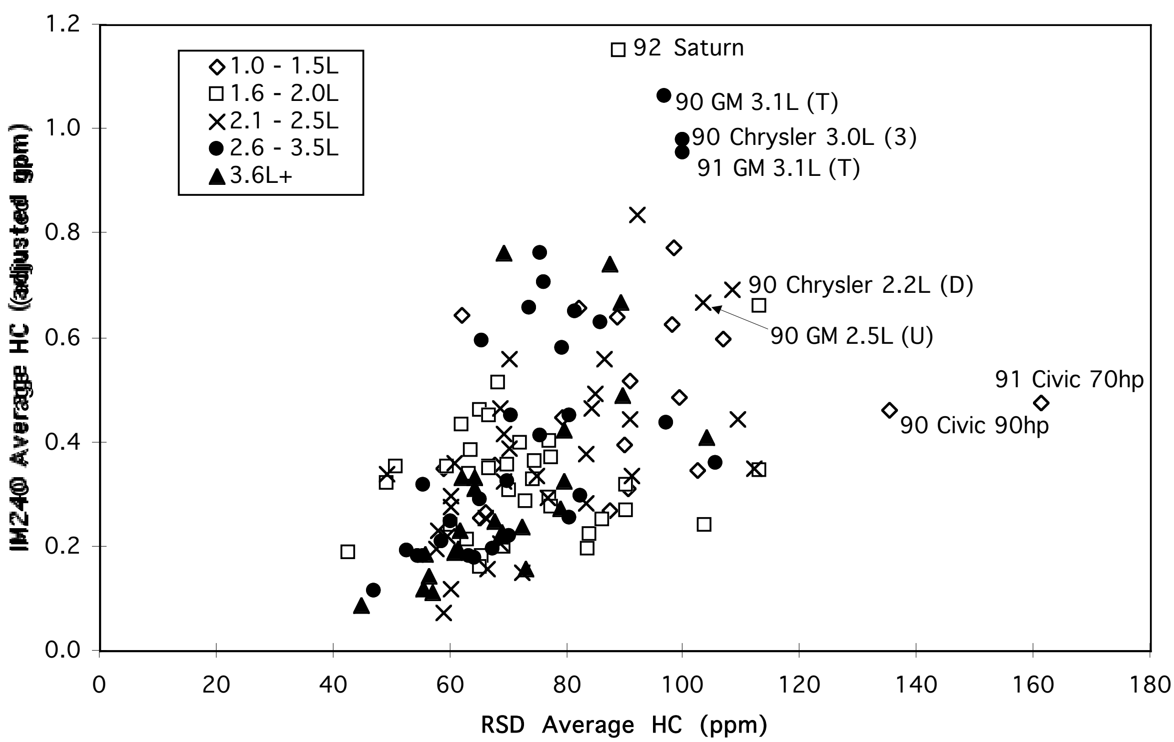


Table 1. IM240 and Remote Sensing Emissions and Failure Rates of Models Identified in Figures

	IM240		RSD	
CO	Avg gpm	Fail Rate	Avg %	Fail Rate
91 GM T 3.1L	16.6	13%	0.67	18%
90 GM T 3.1L	16.1	6%	0.61	16%
90 Nissan Sentra	15.6	5%	0.87	26%
91 GM E 5.0L	12.5	11%	0.74	21%
90 Honda Civic 70hp	8.7	0%	1.06	34%
90 Honda Civic 92hp	8.1	0%	1.03	33%

	IM240		RSD	
HC	Avg gpm	Fail Rate	Avg ppm	Fail Rate
92 Saturn 1.9L MFI	1.15	29%	89	10%
90 GM T 3.1L	1.07	6%	97	10%
91 GM T 3.1L	0.96	13%	100	9%
90 Chrysler 3 3.0L	0.98	4%	100	12%
90 Chrysler D 2.2L	0.69	4%	108	14%
90 GM U 2.5L	0.67	3%	103	14%
91 Honda Civic 70hp	0.48	1%	161	13%
90 Honda Civic 92hp	0.46	0%	135	11%

APPENDIX E

DRAFT LBL-41451

I/M FAILURE RATES BY VEHICLE MODEL

Tom Wenzel, Lawrence Berkeley National Laboratory
Marc Ross, University of Michigan

Presented at the
Seventh Coordinating Research Council On-Road Vehicle Emissions Workshop
San Diego, CA
April 9-11, 1997

Energy Analysis Program
Environmental Energy Technologies Division
Lawrence Berkeley National Laboratory
University of California
Berkeley, CA 94720

This work was supported by the Office of Mobile Sources, Office of Air and Radiation, of the U.S. Environmental Protection Agency. Prepared for the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

Abstract

Previous analysis of vehicle emissions measured by roadside infrared sensors in Califonia indicates that vehicle emissions vary greatly by vehicle model. This analysis compares vehicle emission data by model year and model from two additional states using different measurement techniques to determine if differences in failure rates by vehicle model are consistent. Particular attention is paid to Arizona IM240 test results on relatively new (2- to 5-years old) passenger car models.

Previous Results from CA RSD Data

In a previous study we examined in-use emissions data on large numbers of vehicles from remote sensing measurements taken in California (Ross et al 1995; Wenzel and Ross 1996a, 1996b). The California Air Resources Board's database includes 90,000 remote sensing measurements of about 60,000 vehicles, taken at several California sites in 1991. We took two steps to ensure that the remote sensing data were giving an accurate picture of in-use emissions, and were not measuring emissions during enrichment events. First, we only considered readings from sites where very few acceleration events were observed. Second, where possible, we averaged multiple readings of the same vehicle.

Figure 1 shows the failure rate (defined as exceeding 1% CO concentration) against the average CO concentration, for all cars of a given MY-model combination from MY87-89. In the figure, each point represents a particular MY-model combination (for example, 1987 Nissan Sentras). CO concentration for a model correlates well with the failure rate of that model. This is because just a few high emitters in a particular model will cause a noticeable shift in the mean for that model. There also is a wide range in failure rate, from no failures to nearly 30 percent failures, depending on the particular model. Remember that these vehicles are 2- to 5-years old, within the manufacturer's warranty period for emission control components. Cars from five relatively inexpensive models of Asian manufacture (open circles) have an average failure rate of 17 percent. The average failure rate of the remaining models is 4 percent. Models from the 3 major Asian manufacturers appear in both groups. For example, 3 model years of Nissan Sentras, which is the least expensive Nissan analyzed, are among the highest emitters, while 3 model years of Nissan Maximas are among the lowest emitters. Most of the cars from these five models are carbureted, although some fuel-injected models are among the worst, and some carbureted models are among the best.

Comparison of Data From Three States

We next compared the California remote sensing data with I/M data from two states; we obtained one year of idle test data from Minnesota and eleven months of IM240 test data from Arizona. Table 1 compares the three types of emission tests. While the California and Minnesota data were collected in the same year, the IM240 data are from 1995: therefore cars of any given model year are 4 years older in the IM240 data. Idle and IM240 testing is more detailed than the RSD measurements: all three pollutants are measured, and testing is conducted over several minutes rather than one second.

However, there are limitations to all three types of data: the vehicle is not tested under load in the idle test, and vehicles are tested over different durations in the IM240 test. It also appears that inconsistent preconditioning is a problem with the IM240 data. For the most part, we study actual failure rates from the IM240 data, based on the interim cutpoints, to eliminate any bias from unequal testing durations and inconsistent preconditioning.

Table 1: Characteristics of Datasets from Three States

	RSD (California)	Idle (Minnesota)	IM240 (Arizona)
Test Year	1991	1991	1995
MY87-89 Vehicles:			
Number of Tests	15,000	409,000	180,000
Time Period	2 months	12 months	11 months
Vehicle Age	2 - 5 years	2 - 5 years	6 - 9 years
Pollutants	CO, HC	CO, HC, NOx	CO, HC, NOx
Cutpoints Used	CO: 1%	CO: 1%	CO: 30, 15 gpm HC: 2.0, 0.8 gpm NOx: 3.0, 2.0 gpm
Limitations	enrichment? cold starts? 1-sec. snapshot	vehicle not under load	different test duration; inconsistent pre-conditioning

Figure 2 shows CO failure rates of the MY87-89 models based on remote sensing and IM240 testing. The open circles are the five worst models identified by remote sensing. The range in failure rates is lower under the IM240 test, even though these vehicles are 4 years older than in the remote sensing data. This is most likely a result of the different types of tests and cutpoints used. The relationship between failure rate and car model appears relatively strong even as these models age.

Figure 3 plots the CO failure rates of the same models, as measured by remote sensing and idle testing. The idle test results in very low failure rates for these models, even when the same cutpoint (1%) is used. In addition, there is little agreement between the remote sensing results and the idle results; models with high failure rates under remote sensing and IM240 tests do not fail the idle test at the same rate. Perhaps this is not surprising, in that remote sensing and IM240 tests measure emissions under varying loaded operating conditions, whereas the idle test does not.

Table 2 shows the failure rates from each type of test, for the five worst models identified by remote sensing. Three of these models have high failure rates under both remote sensing and IM240 testing, particularly when final IM240 cutpoints are used. However, the other two models appear to be relatively clean based on the IM240 results.

The IM240 data identified several additional MY87-89 domestic models that have high failure rates, as shown in Table 3. We did not find these models in our analysis of the California remote sensing data because there are relatively few of them in California. These domestic models also may have had relatively high failure rates when they were 2- to 5-years old**.**

Results from AZ IM240 Program

We next calculated failure rates for 200 MY91-93 car models for which at least 100 cars were tested Figures 4 through 6 show the actual failure rate, based on Arizona's interim cutpoints (on the x axis) and the implied failure rate based on final cutpoints (on the y axis). The dashed line shows one-to-one correlation between failure rates under the two cutpoints. As expected, the implied failure rate, based on final cutpoints, would be higher than the actual failure rate. The implied final failure rates also correlate well with the actual failure rates.

Table 2. I/M Failure Rates for Five Worst Models as Identified by Remote Sensing

				CARB RSD		AZ IM240		
				all	road sites	start-up	final	MN Idle
MY	Man.	Model	Fuel	>1%	>1%	>30 gpm	>15 gpm	>1%
RSD and IM240 results agree for these 3 foreign models:								
1987	Foreign	A	Carb	30%	26%	17%	26%	10%
1988			Carb	27%	20%	14%	21%	2%
1989			Carb	28%	20%	9%	13%	1%
1987	Foreign	B	Both	27%	20%	11%	15%	1%
1988			FI	23%	14%	6%	12%	4%
1989			FI	22%	16%	4%	15%	0%
1987	Foreign	C	Carb	22%	8%	8%	17%	5%
1988			Carb	14%	14%	4%	8%	2%
1989			Carb	15%	9%	4%	7%	1%
... but not for these 2 foreign models:								
1987	Foreign	D	Carb	9%	4%	1%	4%	1%
1988			FI	13%	10%	1%	2%	1%
1989			FI	28%	26%	0%	2%	2%
1987	Foreign	E	Carb	18%	11%	6%	10%	1%
1988			Carb	18%	18%	2%	6%	0%
1989			Carb	15%	16%	1%	3%	2%
All Other Models				6%	4%	1%	3%	1%

Table 3. I/M Failure Rates of Models not Identified by California Remote Sensing

				CARB RSD		AZ IM240		
				all sites	road sites	start-up	final	MN Idle
MY	Man.	Model	Fuel	>1%	>1%	>30 gpm	>15 gpm	>1%
Domestic models identified by IM240								
1987	Domestic	A	EFI	13%	13%	14%	17%	8%
1988			EFI			7%	10%	4%
1989			EFI			1%	2%	3%
1987	Domestic	B				14%	24%	7%
1987	Domestic	C	Carb			20%	23%	16%
1988			Carb			13%	14%	18%
1989			Carb			5%	6%	11%
1987	Domestic	D	Carb			13%	17%	0%
1987	Domestic	E	Carb			14%	19%	2%
1988			Carb			9%	13%	2%
1989			Carb			7%	11%	0%
1987	Domestic	F	Carb			22%	44%	6%
1988			Carb			16%	9%	5%
All Other Models						2%	4%	2%

Again, we found that a few models (indicated by open circles) have failure rates several times that of all other models. We found 6 models with a combined CO failure rate of 5 percent (9 percent using the final cutpoints), 10 models with a combined HC failure rate of 5 percent (14 percent using final cutpoints), and 7 models with a combined NOx failure rate of 10 percent (14 percent using final cutpoints). Some models had high failure rates for more than one pollutant, but none were among the highest emitters for all three pollutants. Most of these high emitting models are of domestic manufacture, although not all three years of each model were among the worst emitters. The MY92 Saturn 1.9L engine with port fuel-injection failed both EPA and CARB in use compliance testing for HC emissions; this model has the third highest HC failure rate in Figure 8 (about 15%). However, the 1991 version of this model has an even higher failure rate under the Arizona IM240 program. The MY91 Saturn passed compliance testing at both EPA and CARB.

Next we looked at interim and final failure rates by mileage groups for the two worst MY91-93 models, as shown in Table 4. For most MY-models, failure rates are low at low mileage, and may increase at higher mileages. However, for in some model years for each of these models, failure rates are high, even under 50,000 miles when the vehicle is still under warranty (shaded boxes). For example, over 20% of MY91 Model B passenger cars with less than 50,000 miles fail for NOx in the Arizona IM240 program.

Table 4: Selected Passenger Car Model Failure Rates by MY and Mileage, 1995 Arizona IM240

			CO		HC		NOx	
			Start-Up	Final	Start-Up	Final	Start-Up	Final
MY-Model	Miles	Count	>20 gpm	>15 gpm	>1.2 gpm	>0.8 gpm	>2.5 gpm	>2 gpm
MY91 A	<25K	221	***6%***	***9%***	***6%***	***16%***	1%	2%
	25-50K	730	***6%***	***15%***	***7%***	***21%***	0%	1%
	50-75K	781	***10%***	***18%***	***9%***	***26%***	1%	2%
	>75K	395	***12%***	***22%***	***15%***	***31%***	1%	3%
MY92 A	<25K	358	2%	4%	***1%***	***8%***	0%	1%
	25-50K	1010	4%	8%	***4%***	***13%***	1%	1%
	50-75K	628	4%	11%	***5%***	***18%***	0%	0%
	>75K	169	8%	12%	***8%***	***22%***	0%	2%
MY93 A	<25K	603	4%	8%	***2%***	***10%***	0%	0%
	25-50K	1221	4%	8%	***1%***	***8%***	0%	0%
	50-75K	215	3%	7%	***4%***	***13%***	0%	0%
	>75K	82	4%	10%	***2%***	***15%***	0%	0%
MY91 B	<25K	96	0%	0%	***1%***	***13%***	***20%***	***30%***
	25-50K	275	0%	0%	***1%***	***12%***	***25%***	***37%***
	50-75K	228	1%	1%	***3%***	***22%***	***40%***	***49%***
	>75K	95	2%	2%	***12%***	***27%***	***44%***	***57%***
MY92 B	<25K	130	0%	0%	0%	2%	2%	3%
	25-50K	388	0%	0%	0%	1%	1%	3%
	50-75K	154	0%	0%	0%	2%	6%	10%
	>75K	37	0%	0%	3%	14%	11%	19%
MY93 B	<25K	305	0%	0%	0%	0%	0%	0%
	25-50K	363	0%	0%	0%	0%	0%	0%
	50-75K	73	0%	0%	0%	1%	0%	0%
	>75K	12	0%	0%	0%	0%	0%	0%

We also looked at MY91-93 light duty trucks. Figure 7 plots NOx interim and final failure rates for trucks. Again, there appears to be a wide range in failure rate by vehicle model. The model with the highest failure rate is the MY92 GM truck engine family that recently failed CARB in-use compliance testing for NOx. However, the other two years of this engine also are high NOx emitters, as are some other models.

Table 5 shows interim and final NOx failure rates by mileage for the two worst MY91-93 models. Again, failure rates tend to increase with mileage for some years. However, NOx failure rates are consistently high for all three years of Model A and two years of Model B (shown in bold italics). Model A is the engine family that failed recall testing; its recall for repair is being challenged by GM.

Table 5: Selected Light-Duty Truck Model Failure Rates by MY and Mileage, 1995 Arizona IM240

			CO		HC		NOx	
			Interim	Final	Interim	Final	Interim	Final
MY-Model	Miles	Count	>60 gpm	>40 gpm	>2.4	>1.6 gpm	>3.0 gpm	>2.5
MY91 A	<25K	178	0%	3%	6%	12%	***15%***	***15%***
	25-50K	490	0%	2%	2%	7%	***8%***	***10%***
	50-75K	767	1%	2%	3%	9%	***12%***	***13%***
	>75K	611	0%	3%	4%	13%	***14%***	***16%***
MY92 A	<25K	174	0%	0%	2%	7%	***20%***	***21%***
	25-50K	688	0%	0%	3%	9%	***24%***	***26%***
	50-75K	501	0%	0%	2%	8%	***30%***	***31%***
	>75K	220	1%	1%	5%	11%	***36%***	***39%***
MY93 A	<25K	333	0%	0%	1%	3%	***14%***	***17%***
	25-50K	669	0%	0%	1%	4%	***20%***	***23%***
	50-75K	225	0%	1%	1%	4%	***23%***	***24%***
	>75K	71	0%	0%	1%	4%	***31%***	***37%***
MY91 B	<25K	39	0%	0%	3%	8%	***3%***	***10%***
	25-50K	132	0%	0%	2%	4%	***9%***	***11%***
	50-75K	219	0%	0%	1%	7%	***13%***	***15%***
	>75K	121	0%	0%	6%	13%	***20%***	***22%***
MY92 B	<25K	59	2%	2%	3%	7%	***14%***	***17%***
	25-50K	153	0%	0%	0%	5%	***10%***	***12%***
	50-75K	90	1%	1%	1%	7%	***19%***	***21%***
	>75K	27	0%	0%	7%	7%	***22%***	***22%***
MY93 B	<25K	88	0%	0%	0%	0%	7%	8%
	25-50K	166	1%	1%	1%	3%	4%	8%
	50-75K	34	0%	0%	0%	3%	3%	9%
	>75K	14	0%	0%	0%	7%	21%	21%

Failure Rate by Station

Finally, we looked at failure rates by test station to determine if aggregate driver differences affect failure rates. We focused on two stations that represent the extremes in average driver income: Station 4, which is in a relatively low-income area, and Station 10, in a relatively high-income area.

Figure 8 shows the average HC failure rate for cars by model year and by test station. Cars tested at Station 4 have significantly higher, and cars tested at Station 10 significantly lower, failure rates than cars tested at all other stations. We suspect this large difference is due to the income differences between the two areas, which may result in inadequate vehicle maintenance in the less affluent neighborhood. (The increase in failure rates in MY91 cars is due to stricter cutpoints applied to MY91 and newer cars).

We next examined model-specific failure rates for MY91-93 cars at each test station. We only considered those models that had at least 100 cars in the entire dataset, and at least 20 cars at each of the two test stations. Figure 9 ranks the 71 models by increasing overall HC failure rate (diamonds). The circles are paired failure rates of each model at the two test stations (closed circles are from Station 4, open circles from Station 10). Models that have no circles had no failures at both test stations. Over half of these MY91-93 models are consistently clean at each station. However, for virtually every other model, the failure rate is higher at Station 4 than at Station 10. Many of the models failing at Station 4 show no failures at Station 10. This finding is even more striking when we look at average HC results by model and test station. These plots suggest that I/M failure rates for modern, 2- to 5-year old cars are sensitive to how individual drivers treat their cars. These plots also indicate that manufacturers can design cars that have essentially no failures within 5 years.

Uses of Failure Rates by Model

We believe that careful analysis of state I/M data can be used for several purposes. First, a list of vehicle models and their failure rates can be published, to give consumers information on the likelihood that their car will fail an I/M test in the future. Published rankings of failure rates by model will also make the relative emission levels of models more prominent, and may spur manufacturers to improve the effectiveness and durability of their emissions control systems.

Second, failure rates by model could be used to target specific engine families for further, more detailed in-use compliance testing under existing EPA and CARB programs. These programs have limited effectiveness in that only a few engine families, and that only a few (10 to 15) vehicles of each of these families, are tested each year. State I/M data from hundreds of thousands of vehicles can identify engine families that are suspected high emitters, and agency resources can be targeted towards testing a truly representative sample of those engines.

Third, the data can be used to evaluate the effectiveness of different I/M programs in different states. If a model with known high failure rates is frequently passed in a particular state, this is an indication that there is a problem in that state's I/M program.

Finally, at the state level, failure rates by vehicle model can be used to identify potential high emitters for more detailed I/M testing, such as the High Emitter Profile California is proposing. Or they can be used to identify potential low emitters to be waived from I/M requirements, as Arizona is considering.

Summary

Our analysis indicates that failure rates and average emissions by vehicle model from idle I/M test data do not correlate with remote sensing measurements or IM240 test results. IM240 data support earlier results for for MY87-89 cars using remote sensing measurements: a few models have failure rates several times that of all other cars; differences in failure rate by model continue as vehicles age, up to 9 years; and IM240 data indicate that a few domestic models may also have been among the worst CO emitters of MY87-89 cars. The IM240 data show similar results for MY91-93 car models. In addition, many models with high average CO emissions also have high average HC emissions, while many models have high HC and NOx emissions. In contrast with our earlier finding, most of the worst MY91-93 models are of domestic manufacture. For most car models, average emissions tend to increase with increasing mileage; however, some models have high average emissions under 50,000 miles. The IM240 data also correctly identify certain engine families that have been recalled under EPA and CARB In-Use Compliance testing programs. Finally, our analysis of emissions by vehicle model and I/M test station suggests that poor vehicle maintenance affects emissions and failure rates, even for new, modern technology cars; however, most modern car models do not appear to be sensitive to poor maintenance

Figure 1. CO Failure Rate and Average Concentration
MY87-89 Car Models, 1991 California Remote Sensing

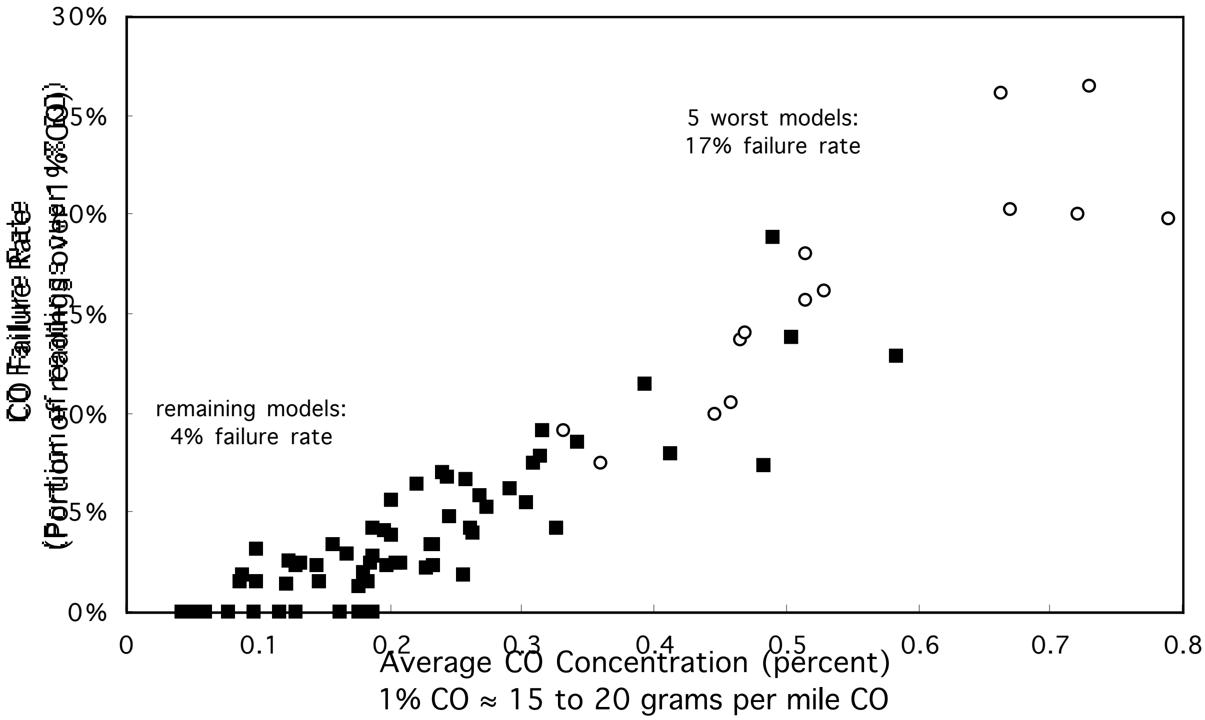


Figure 2. CA RSD and AZ IM240 CO Failure Rates by Model
MY87-89 Models

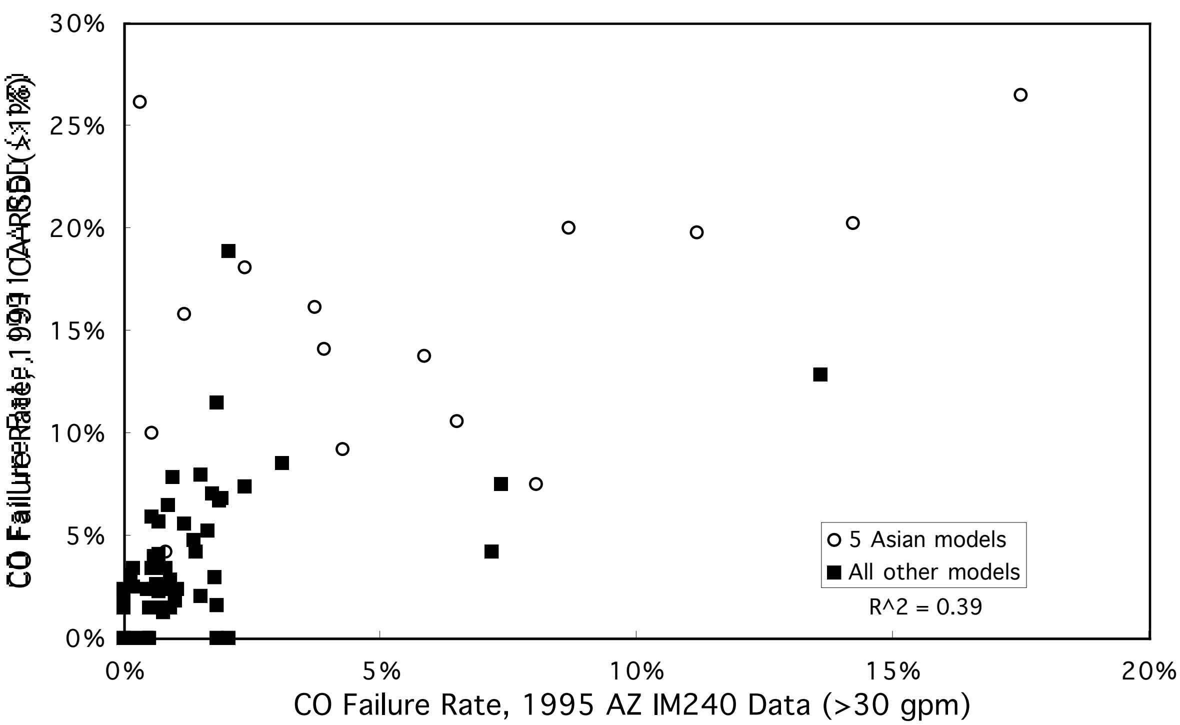


Figure 3. CA RSD and MN Idle CO Failure Rates by Model
MY87-89 Models

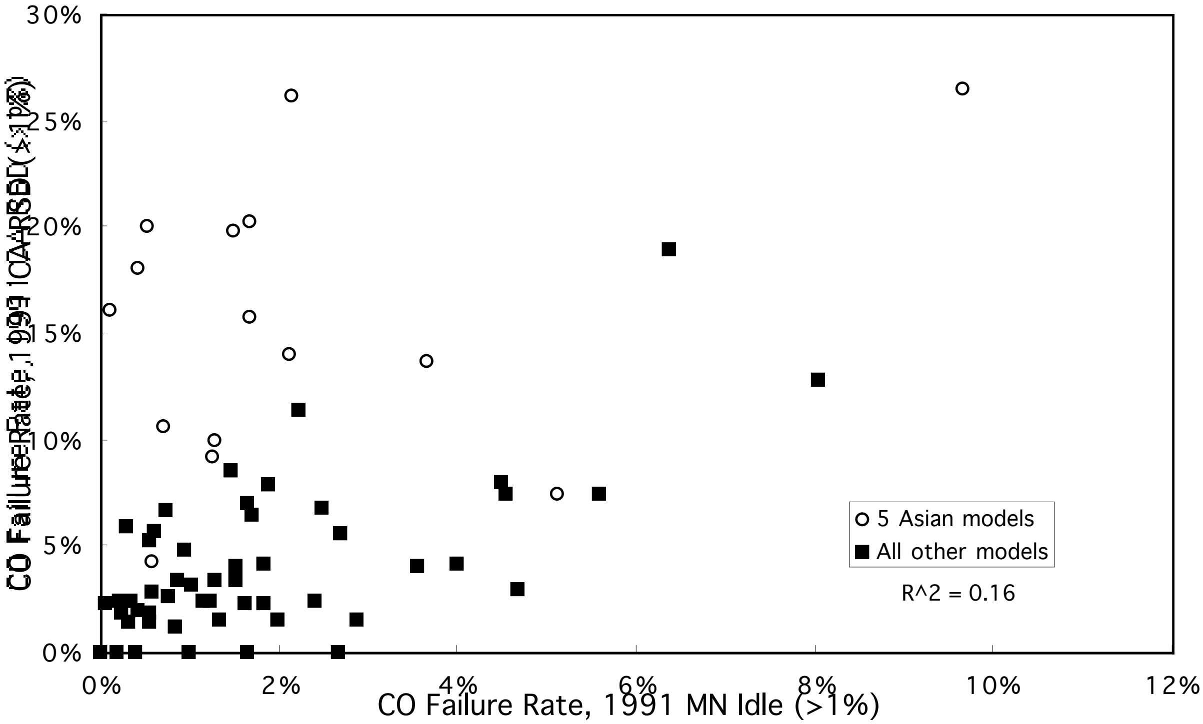


Figure 4. CO Failure Rates, MY91-93 Car Models
1995 AZ IM240 Data

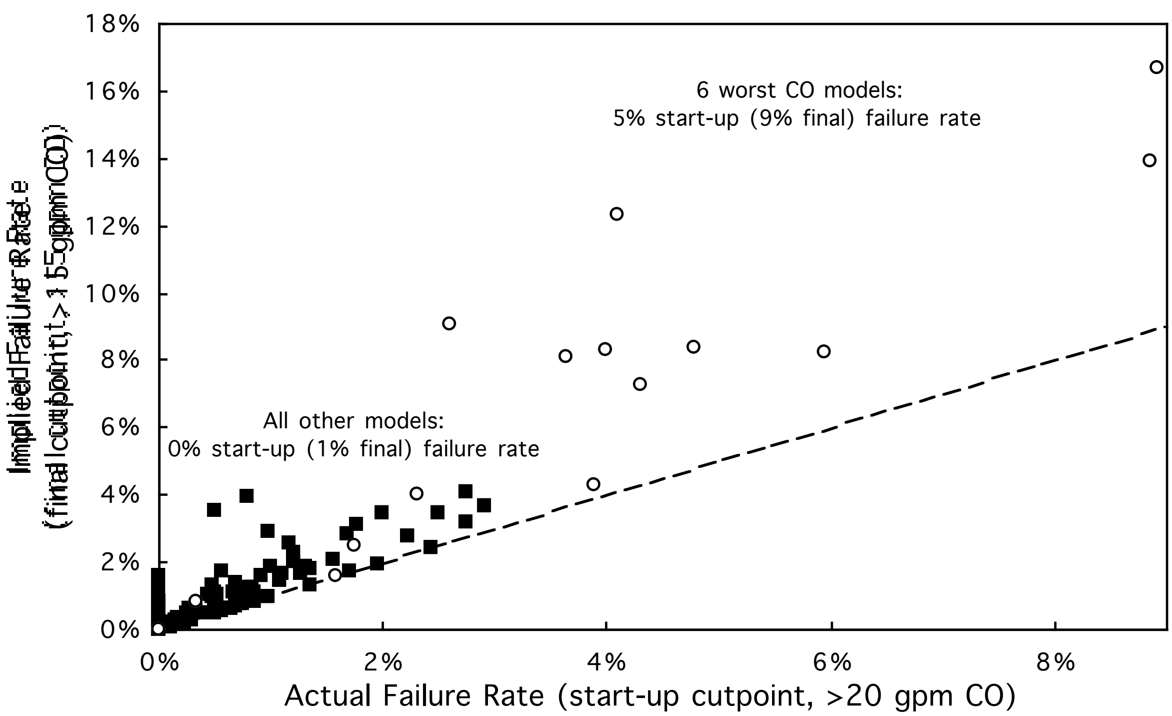


Figure 5. HC Failure Rates, MY91-93 Car Models

1995 AZ IM240 Data

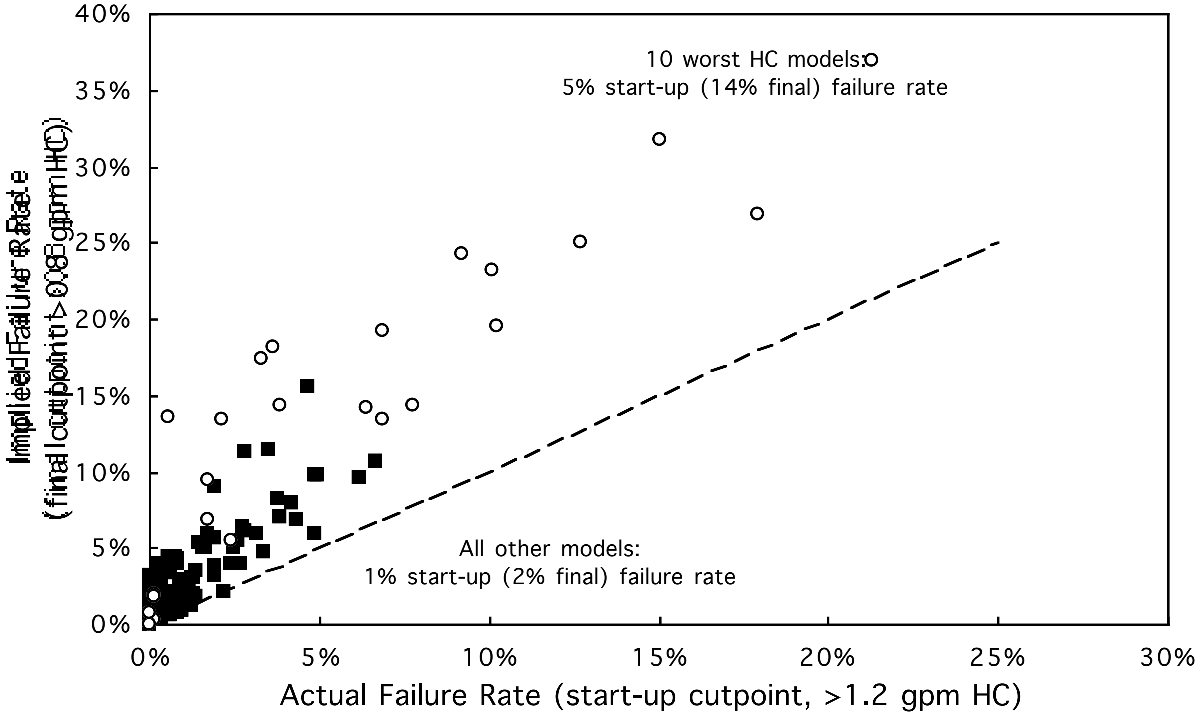
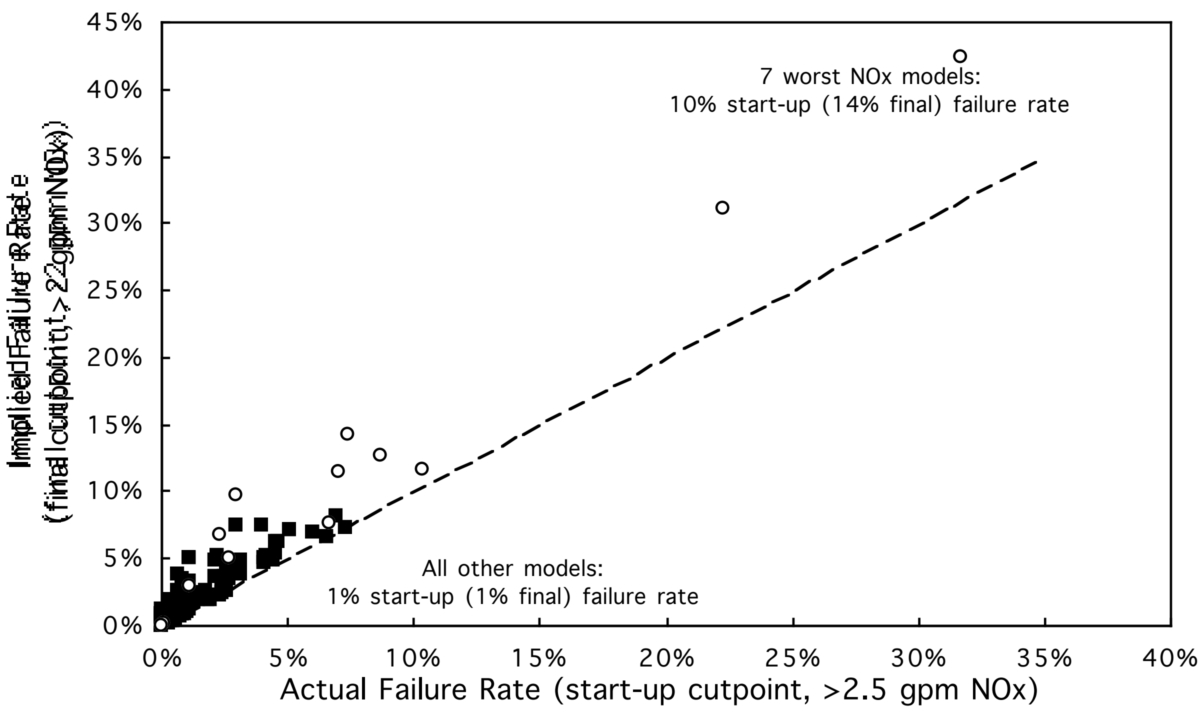


Figure 6. NOx Failure Rates, MY91-93 Car Models

1995 AZ IM240 Data



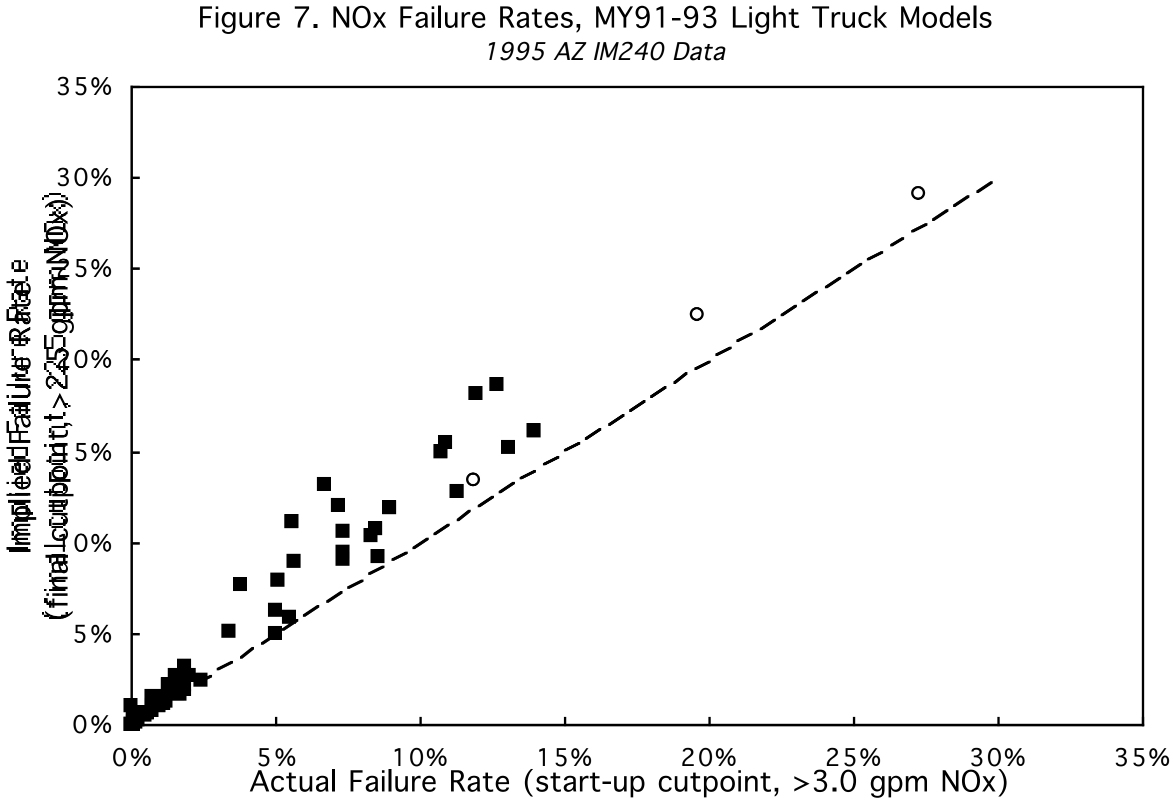


Figure 7. NOx Failure Rates, MY91-93 Light Truck Models
1995 AZ IM240 Data
Implied Failure Rate (final cutpoint, >2.5 gpm NOx)
Actual Failure Rate (start-up cutpoint, >3.0 gpm NOx)
0%
5%
10%
15%
20%
25%
30%
35%

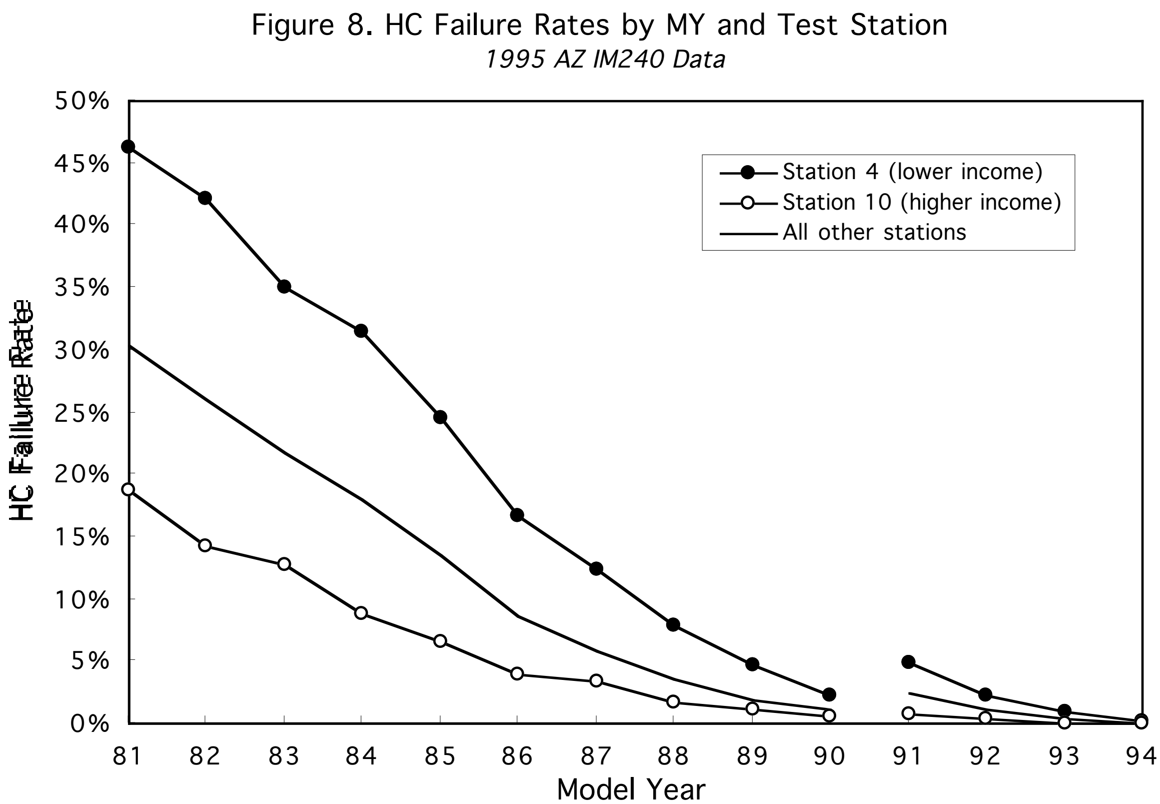


Figure 8. HC Failure Rates by MY and Test Station
1995 AZ IM240 Data
Station 4 (lower income)
Station 10 (higher income)
All other stations
HC Failure Rate
Model Year
0%
5%
10%
15%
20%
25%
30%
35%
40%
45%
50%
81
82
83
84
85
86
87
88
89
90
91
92
93
94

Figure 9. HC Failure Rate by Test Station, MY91-93 Car Models
1995 AZ IM240 Data

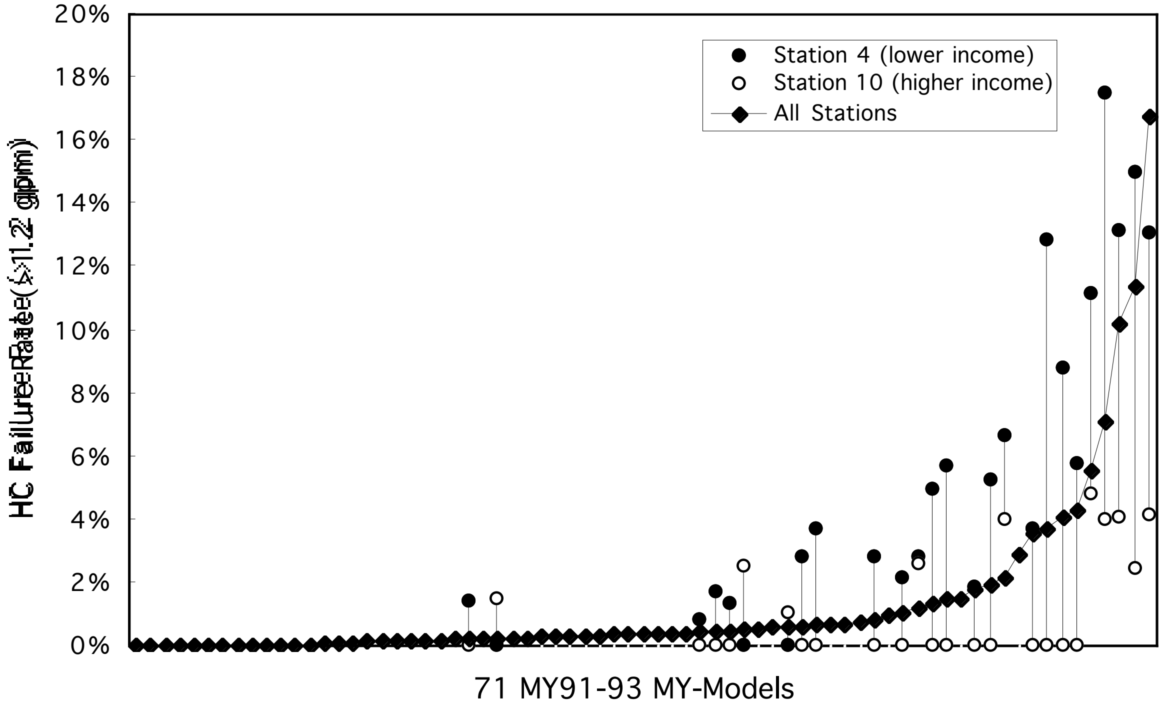
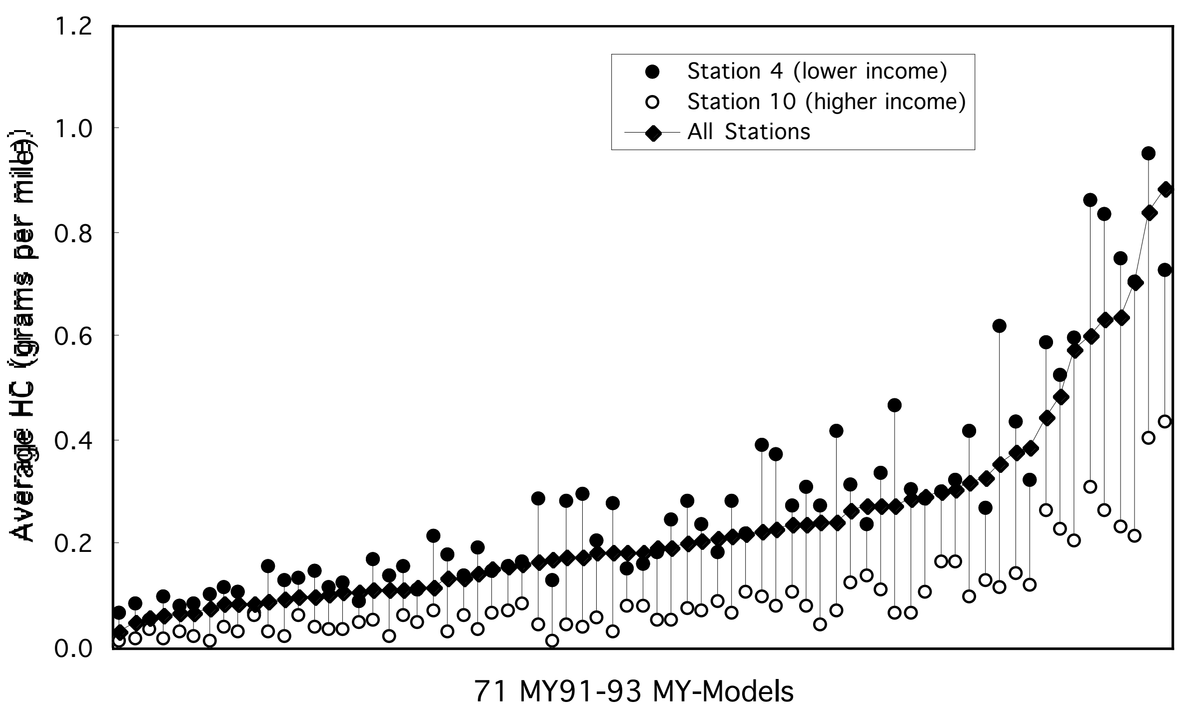


Figure 10. Average HC by Test Station, MY91-93 Car Models
1995 AZ IM240 Data



References

Ross, Marc, Rob Goodwin, Rick Watkins, Michael Q. Wang and Tom Wenzel. 1995. *Real-World Emissions from Model Year 1993, 2000, and 2010 Passenger Cars*. Lawrence Berkeley National Laboratory. LBL-37977. November.

Wenzel, Tom and Marc Ross. 1996a. *Emissions from Modern Passenger Cars with Malfunctioning Emissions Controls*. Society of Automotive Engineers Technical Paper Series. SAE 960067.

Wenzel, Tom and Marc Ross. 1996b. "Relationship between Vehicle Attributes and Malfunction of Emissions Control Systems." Proceedings of the Sixth CRC On-Road Vehicle Emissions Workshop. March.

APPENDIX F

Further Analysis of I/M Failure Rates by Vehicle Model

Tom Wenzel, Lawrence Berkeley National Laboratory

Poster presented at the Eighth CRC On-Road Vehicle Emissions Workshop, April 20-22, 1998, San Diego CA.

Last year we presented an analysis of IM240 failure rates by vehicle model from the Arizona IM240 testing program in 1995. This year we look at three test years of Arizona data to examine the internal consistency of the data. We also present a comparison of failure rates and average emissions by vehicle model from several IM240 programs.

Consistency of AZ data

We compared IM240 failure rates by model for three test years, 1995 through 1997. We only include the first six months of each year in the analysis, in order to ensure that at least 6 months elapsed between all of the tests conducted in each test year. Arizona operates a biennial inspection program, with half of all vehicles from each model year tested every year. Vehicles are not tested when they are resold. Therefore, the same vehicles that were tested in 1995 are tested again in 1997, unless they moved out of the I/M area or were scrapped.

Figure 1 plots HC failure rates for 194 model year 1990 to 1993 car models tested in 1996 against those tested in 1997. Each point on the figure represents a model year/model (e.g. 1993 Nissan Sentra), for which at least 80 individual cars were tested. The failure rates are quite similar in the two test years, with the same models having the highest failure rates in each year. The figure is typical of the other year-to-year comparisons for each pollutant. The table shows that the r-squared values for each year-to-year comparison for each pollutant range from 0.66 to 0.90. The comparisons with the lowest (Figure 2) and highest (Figure 3) correlation are also presented.

Table 1. Consistency of AZ IM240 Model Failure Rates over Time: R^2 of Year-to-Year Comparisons

Test Year Comparison	HC	CO	NOx
194 MY90-93 Car Models			
1995 vs. 1996	0.84	0.74	0.90
1996 vs. 1997	0.78	0.70	0.88
1995 vs. 1997	0.76	0.66	0.80
154 MY87-89 Car Models			
1995 vs. 1996	0.87	0.89	0.85
1996 vs. 1997	0.86	0.87	0.82
1995 vs. 1997	0.85	0.87	0.77

Close analysis of the plots indicates that CO failure rates by MY/model tend to increase over time. That is, the points in Figure 4 tend to be above the (solid) 45-degree line, indicating that CO failure rates by model increased from 1995 to 1996. This is to be expected, since the individual vehicles tested in 1996 would on average be at least 6 months older, and likely have accumulated more mileage, than the vehicles tested in 1995. We see a similar increase in failure rate from 1995 to 1997 (Figure 2); these are the same cars that went through the I/M program in 1995. It appears that the Arizona I/M program roughly offsets one year of the "natural" emissions increase due to vehicle aging and mileage accumulation. We are further analyzing the effect of the I/M program on failure rates and emissions by tracking individual vehicles over several test years.

Figure 1. HC Failure Rates, 194 MY90-93 Car Models, AZ IM240
(at least 80 individual cars tested for each model)

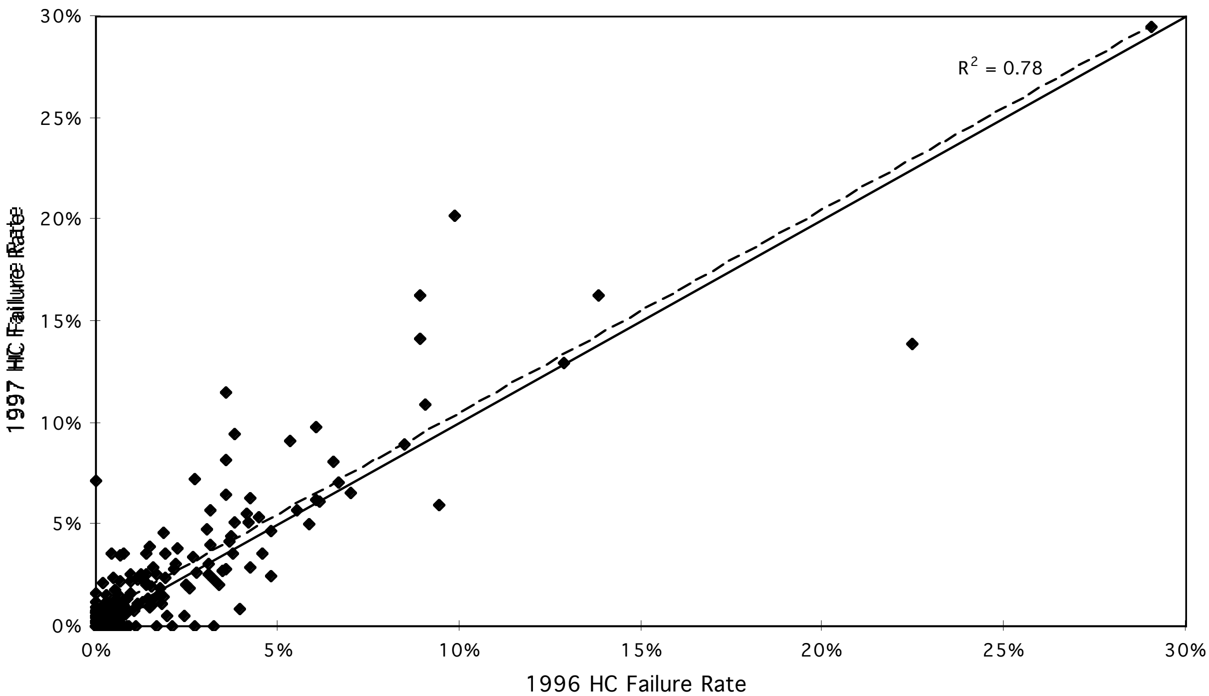


Figure 2. CO Failure Rates, 194 MY90-93 Car Models, AZ IM240
(at least 80 individual cars tested for each model)

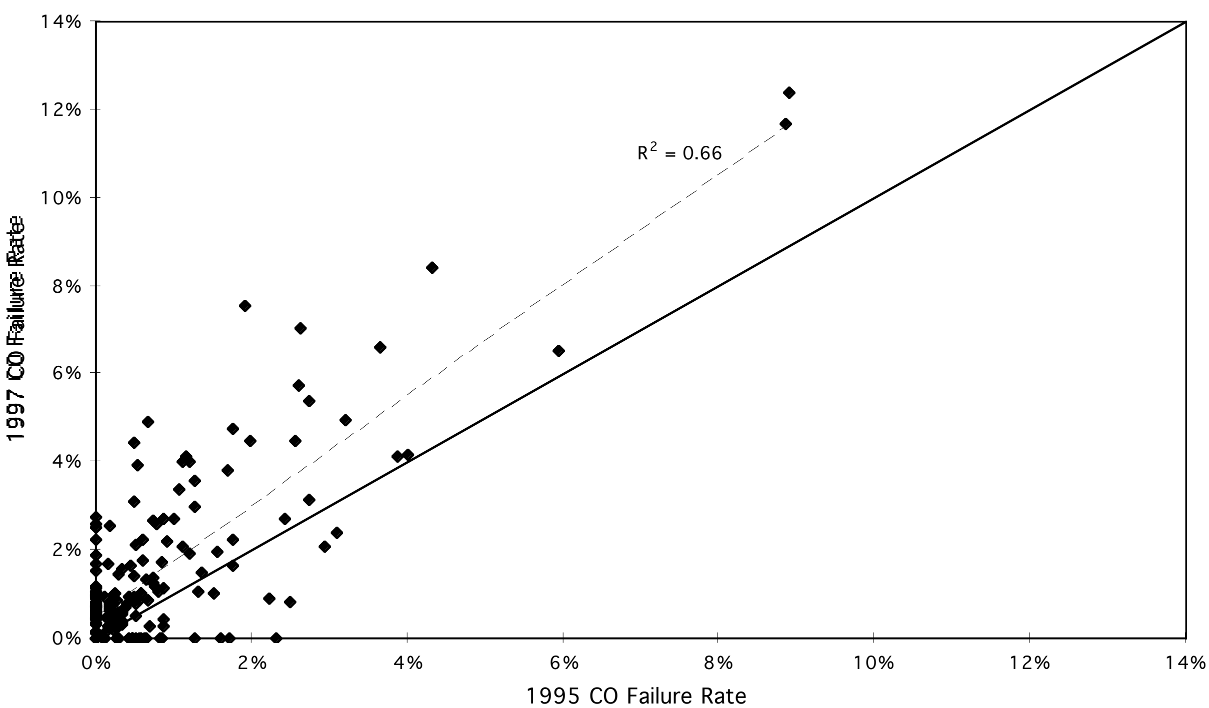


Figure 3. NOx Failure Rates, 194 MY90-93 Car Models, AZ IM240

(at leat 80 individual cars tested for each model)

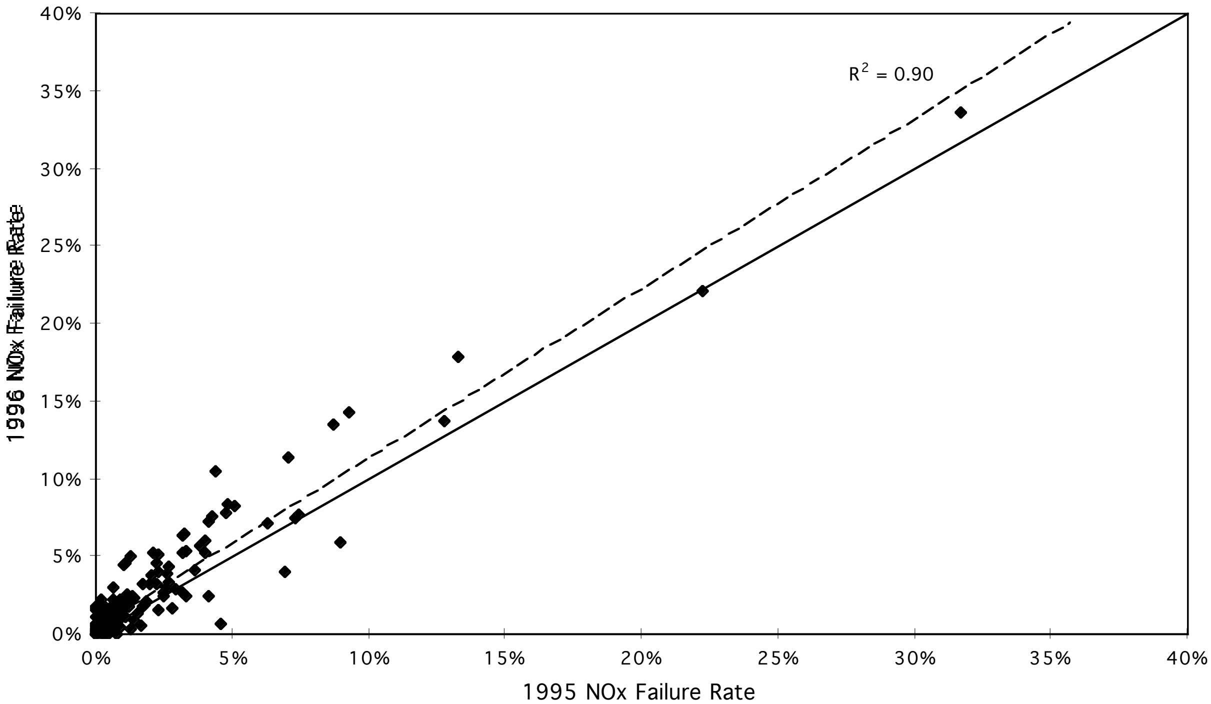
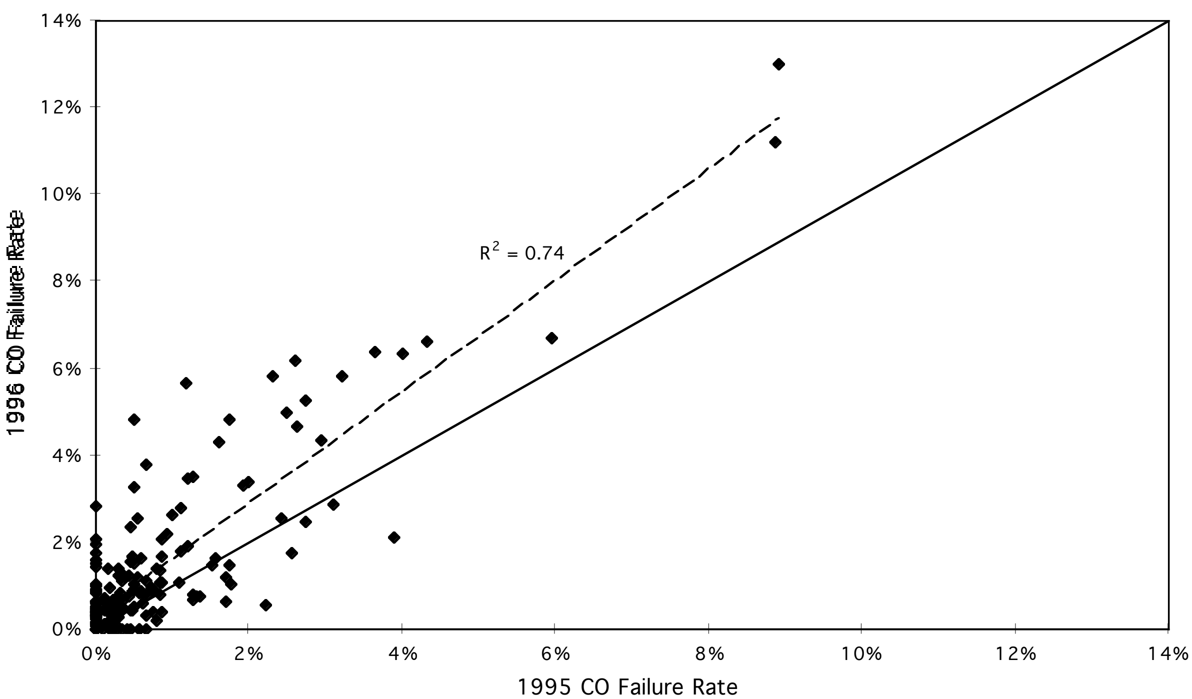


Figure 4. CO Failure Rates, 194 MY90-93 Car Models, AZ IM240

(at least 80 individual cars tested for each model)



Accuracy of I/M emissions results

We next compared failure rates by model to average emissions by model. A limitation of I/M data is that testing procedures are not consistent between vehicles. States like Arizona allow the cleanest vehicles to pass after 30 seconds of testing (fast passes), while the dirtiest vehicles are failed after 94 seconds (fast fails). This means that all vehicles are not tested over the same portion of the IM240 test procedure. To compare emissions from vehicles tested over different portions of the IM240, one needs to correct fast pass/fast fail emissions to full test equivalent values. We use a simple methodology to convert short test results to full test equivalent emissions. This methodology uses correction factors based on the average ratio of emissions at each second to full test emissions, for each pollutant and second of the IM240. For our purposes here, we do not require that this correction results in absolute accuracy for individual vehicles; rather we look for consistent ranking of models.

Arizona runs full IM240 tests on a random sample of 2 percent of the vehicles in the fleet. (Unfortunately we could not distinguish between vehicles given the full test because they were part of this random sample or because they had emissions close enough to the cutpoints that they could neither fast pass or fast fail. It appears that only one-half of the vehicles given an initial full IM240 were part of the random sample). Figures 5a through 7a compare the average emissions of full tests with those of short tests, by vehicle model, for each pollutant. Each plot shows relatively good agreement between the average emissions of vehicles given the full test, and those either fast passed or fast failed, with r-squared values of 0.68 (for NOx) to 0.87 (for CO). The vehicles given the full test appear to have consistently higher emissions than those fast passed or fast failed. Much of this difference is likely due to the crudeness of our adjustment factors; our analysis indicates that our factors tend to underestimate emissions from all cars, and in particular the cleanest cars which make up the majority of the fleet. It is also possible that the random sample of vehicles receiving the full test are not representative of the fleet. If one is concerned only with relative emissions values, it appears that the IM240 short test emissions values can be used to compare groups of vehicles. If one is interested in absolute emissions, however, a better method to project full test emissions is needed.

Figures 5b through 7b show the value of using the average emissions values for vehicles receiving the short test. These figures are identical to Figures 5a through 7a, with the standard error of each estimate included. The vertical “whiskers” are the error associated with the full test cars, while the horizontal whiskers are the error of the short test cars. As the figures show, an order of magnitude increase in the sample size greatly reduces the uncertainty of the estimate of average emissions by model.

Figure 5a. Average HC by Model, Fast Pass/Fast Fail vs. Full Tests
MY90-93 models with at least 10 full IM240 tests, 1995 AZ IM240

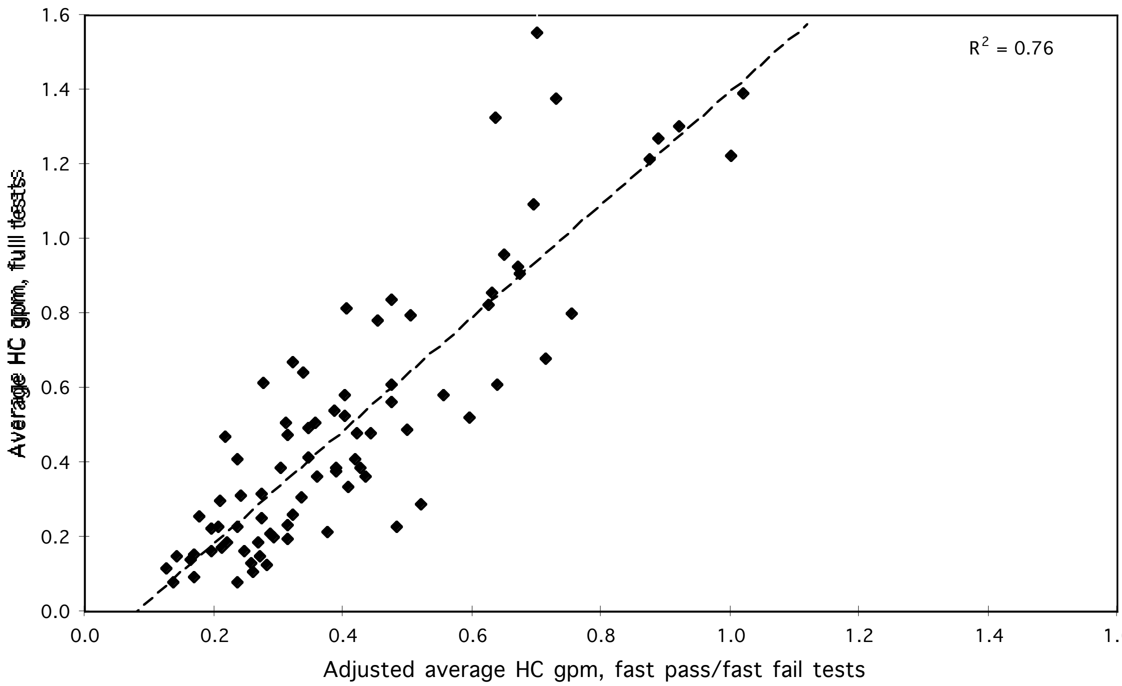


Figure 5b. Average HC by Model, Fast Pass/Fast Fail vs. Full Tests
MY90-93 models with at least 10 full IM240 tests, 1995 AZ IM240

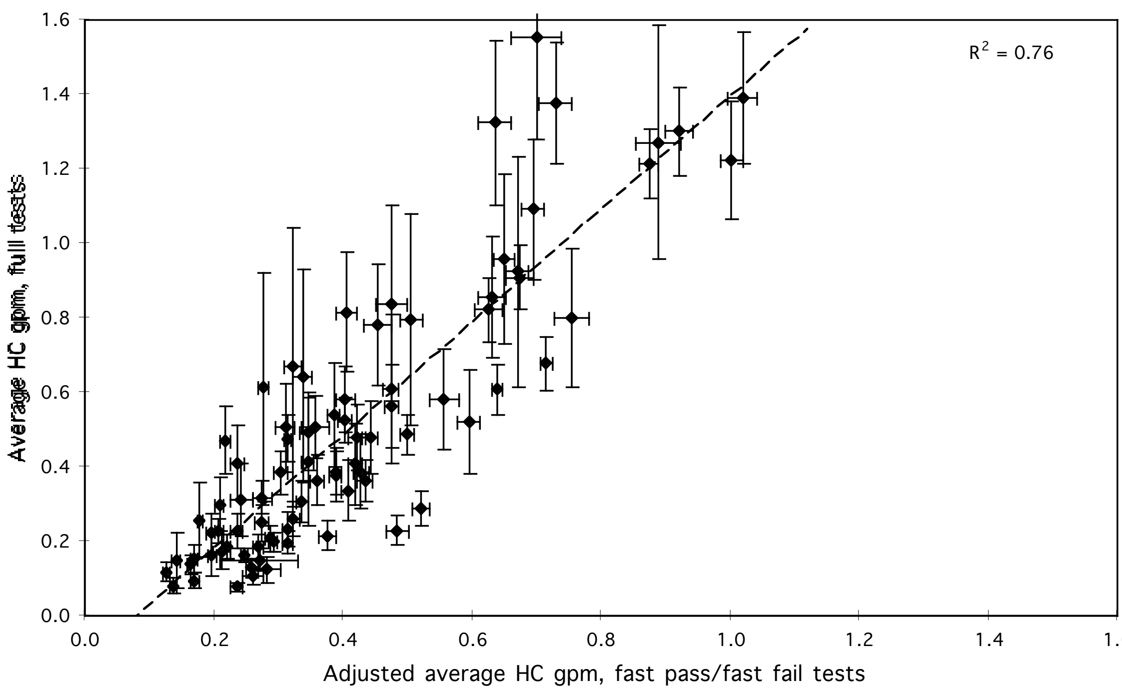


Figure 6a. Average CO by Model, Fast Pass/Fast Fail vs. Full Tests
MY90-93 models with at least 10 full IM240 tests, 1995 AZ IM240

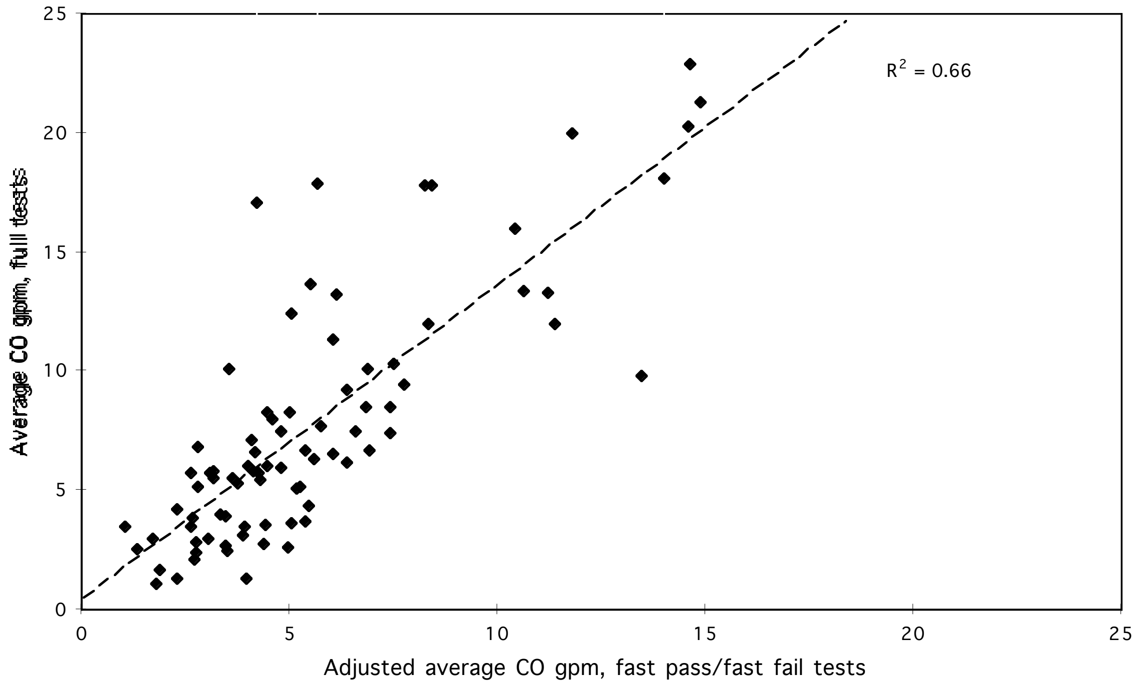


Figure 6b. Average CO by Model, Fast Pass/Fast Fail vs. Full Tests
MY90-93 models with at least 10 full IM240 tests, 1995 AZ IM240

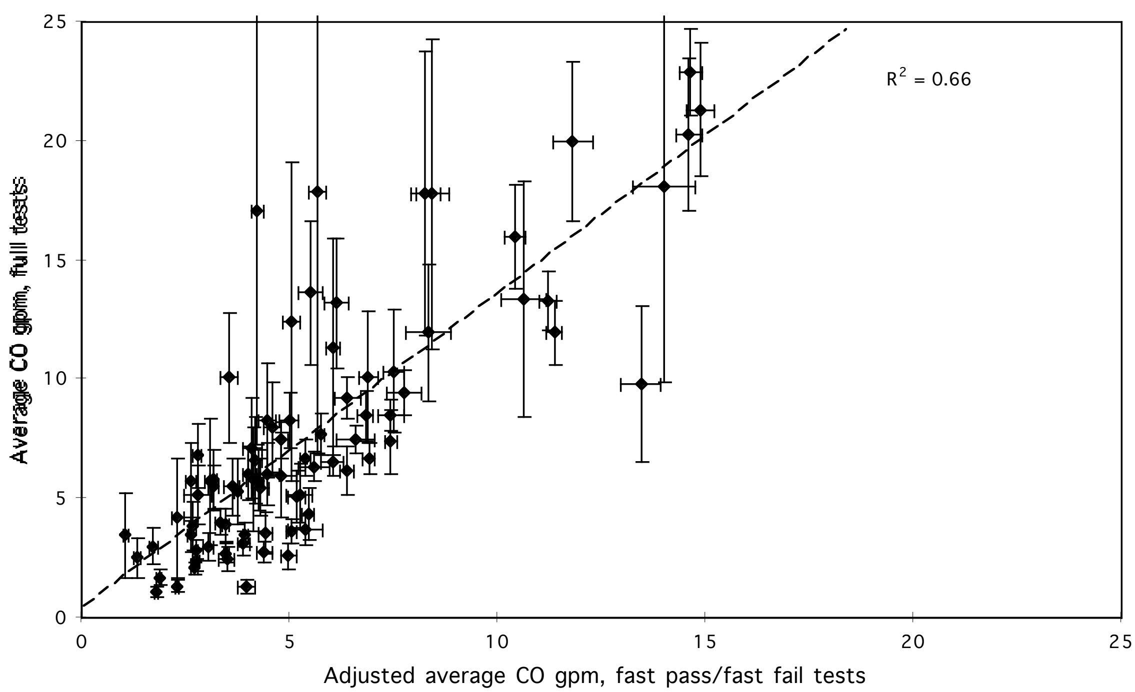


Figure 7a. Average NOx by Model, Fast Pass/Fast Fail vs. Full Tests

MY90-93 models with at least 10 full IM240 tests, 1995 AZ IM240

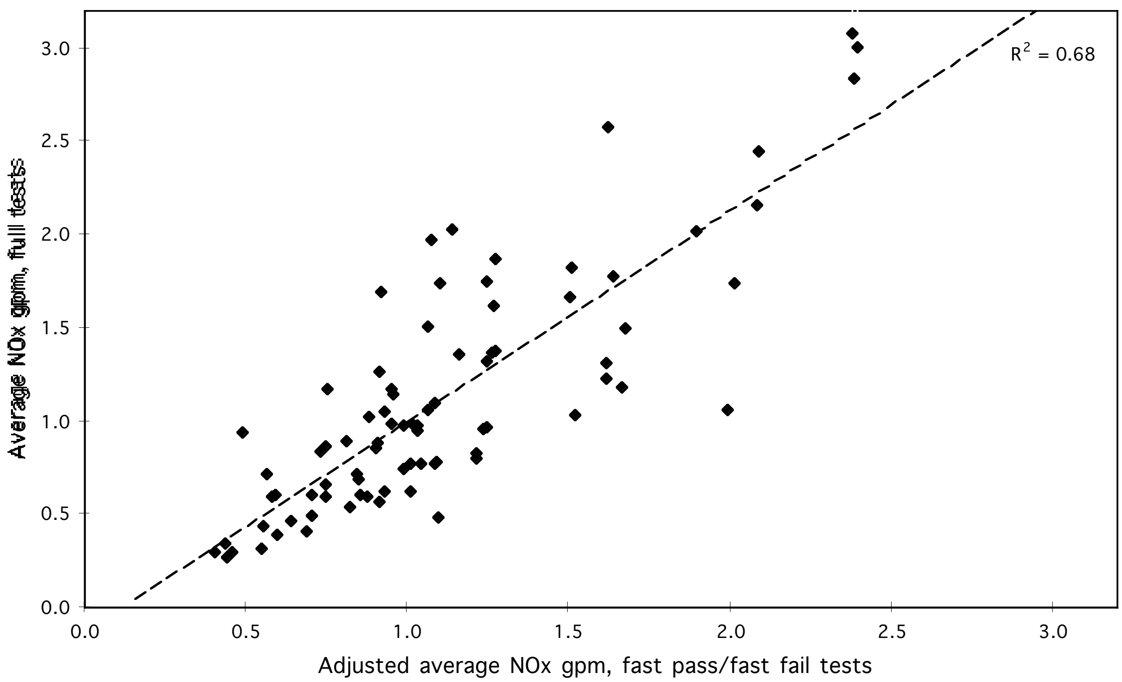
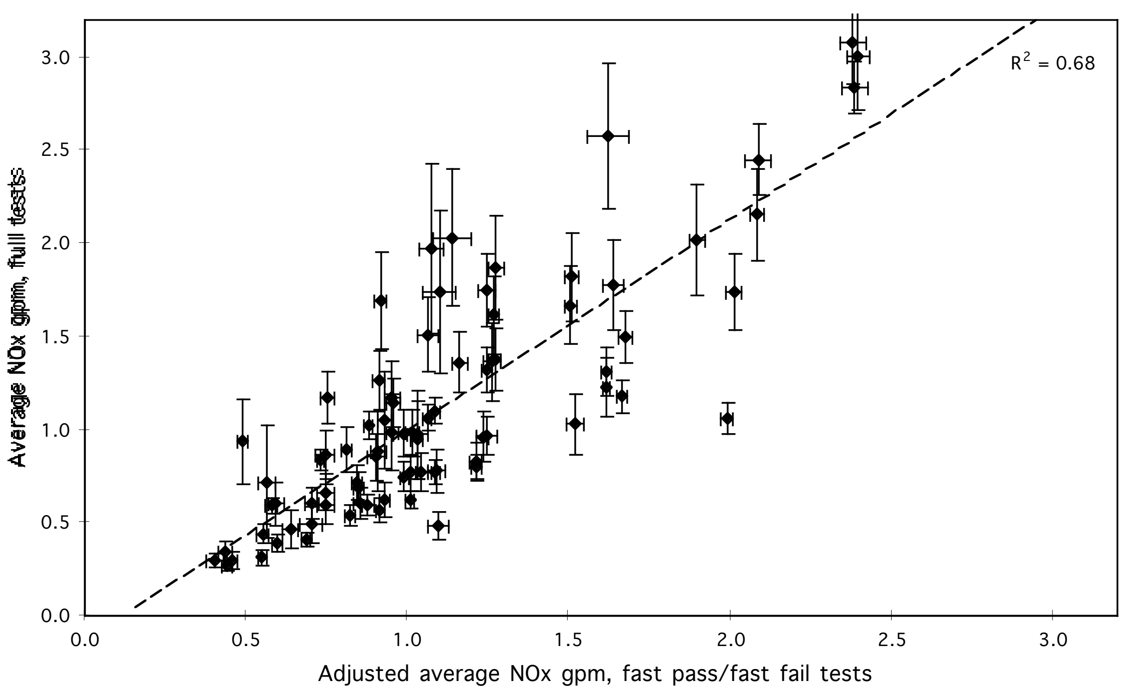


Figure 7b. Average NOx by Model, Fast Pass/Fast Fail vs. Full Tests

MY90-93 models with at least 10 full IM240 tests, 1995 AZ IM240



Interstate comparison of IM240 failure rates

We compared failure rates by model from three states that are using the IM240 test procedure: Arizona, Colorado, and Wisconsin. Table 2 summarizes the key features of each program. Important differences are the cutpoints (Arizona's and Wisconsin's are similar, while Colorado's tend to be less stringent), and the model years tested in each year (while Arizona tests all model years each year, Colorado tested only odd model years in 1997 and Wisconsin tested only even model years). Figures 8 through 10 compare the 1997 combined CO and HC failure rates for 36 selected models; we show CO and HC failure rates only because Wisconsin did not fail vehicles based on NOx emissions. The models studied were chosen because they are popular models and had few engine and transmission options; we wanted to avoid including cars with different sized engines in the same model.

Figure 8 shows good agreement between the Arizona and Wisconsin programs; the relative rankings of the models by failure rate are quite similar. The good agreement may in part be due to the sample of models studied; we only looked at a few models from each of a wide range of model years. Since one expects failure rates to increase as vehicles age, it is no surprise that the oldest models have the highest failure rate in each state. Still, the agreement is quite good if one accounts for model age. The Colorado data do not agree as well with those of the other two states (Figures 9 and 10). This may be due to the different cutpoints used in Colorado. We plan to do a more thorough interstate analysis of failure rate by model, to more fully examine the consistency of emissions by model across different I/M programs.

Table 2. IM240 Program Elements in Three States

Program Element	Arizona	Colorado	Wisconsin (1)
Test Cycle	biennial; all MYs tested in 1997	biennial; odd MYs tested in 1997	biennial; even MYs tested in 1997
Test on Resale?	no	yes	yes
Composite Cutpoints (cars)			
HC	91-95: 1.2 81-90: 2.0	86-95: 4.0 82-85: 5.0	91-95: 1.25 81-90: 2.0
CO	91-95: 20 83-90: 30 81-82: 60	91-95: 20 85-90: 25 83-84: 50 82: 65	91-95: 20 83-90: 30 81-82: 60
NOx	91-95: 2.5 81-90: 3.0	95: 4.0 86-94: 6.0 82-85: 8.0	*91-95: 2.5* *81-90: 3.0*
Fast Pass?	yes	yes	yes
Fast Fail?	yes	no	no
Phase 2 Pass?	yes	no	yes
Second Chance to Pass?	no	yes if emissions <2x cutpoint	yes if emissions <2x cutpoint
Full Tests	random 2%	all vehicles tested 1/97 to 3/97	random 2%

(1) Cutpoints shown were effective 12/96 to 11/97. Although Wisconsin tests for NOx, vehicles are not failed for exceeding NOx cutpoints. Vehicles tested during weekends in 1996 were given full test; this practice was replaced by 2% random sampling in 1997.

Figure 8. AZ v. WI 1997 IM240 Combined CO and HC Failure Rates
Selected Models

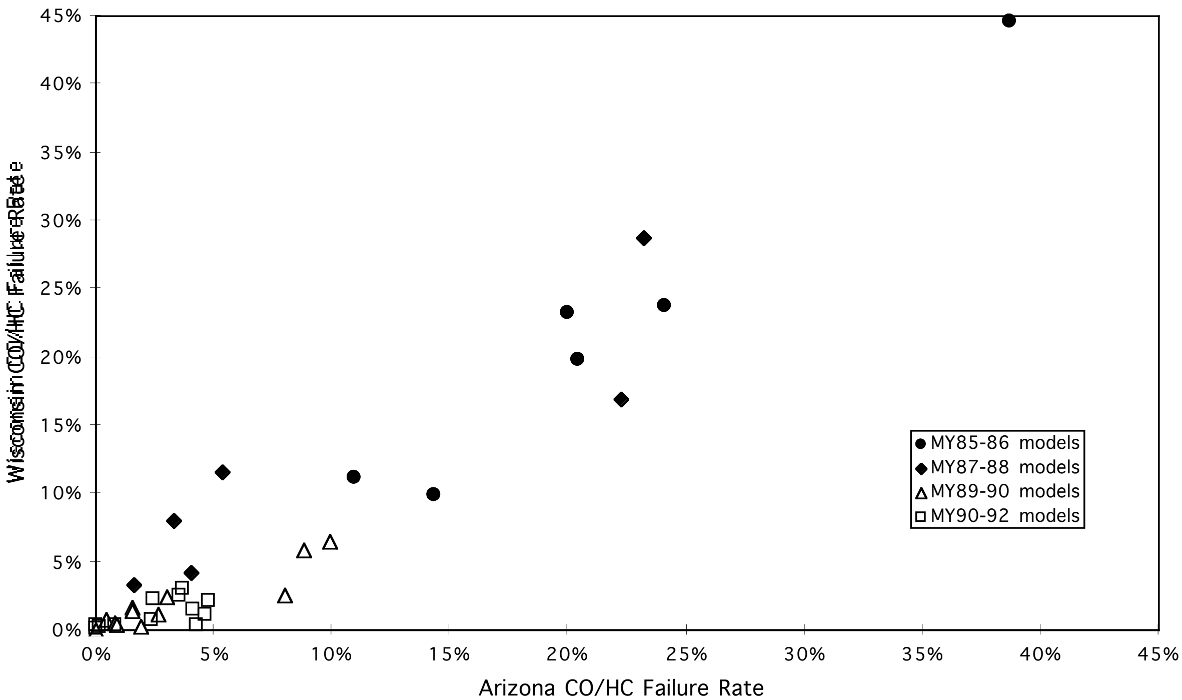


Figure 9. AZ v. CO 1997 IM240 Combined CO and HC Failure Rates,
Selected Models

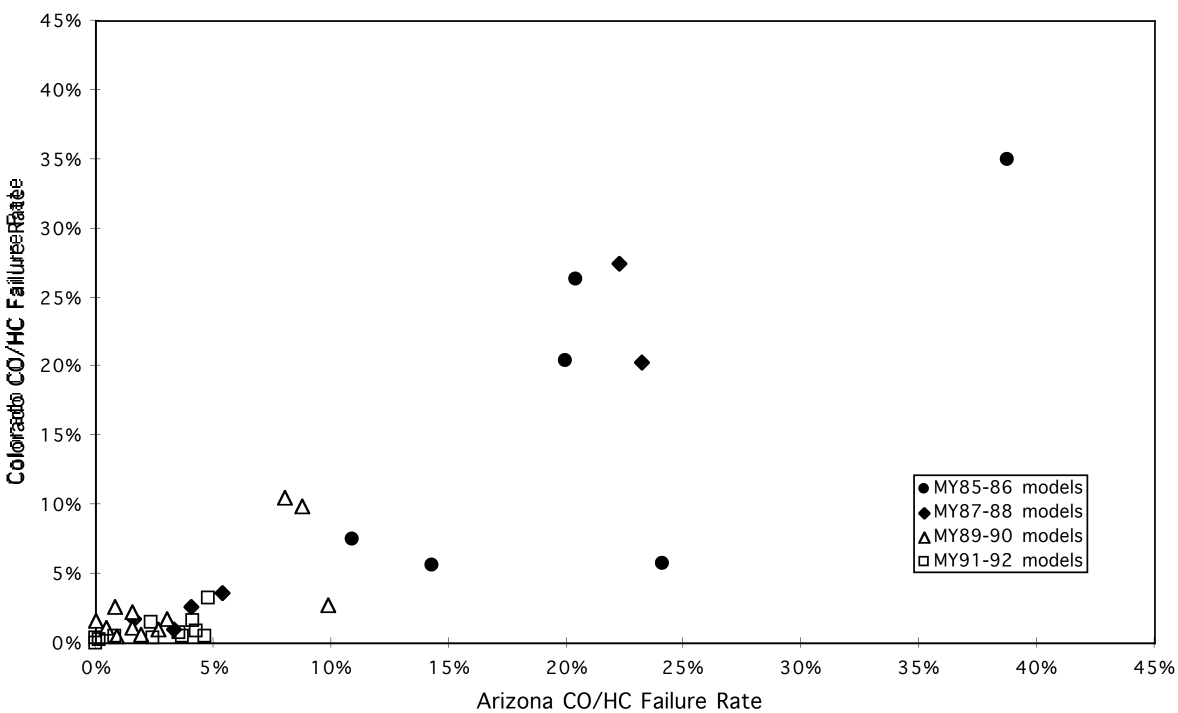
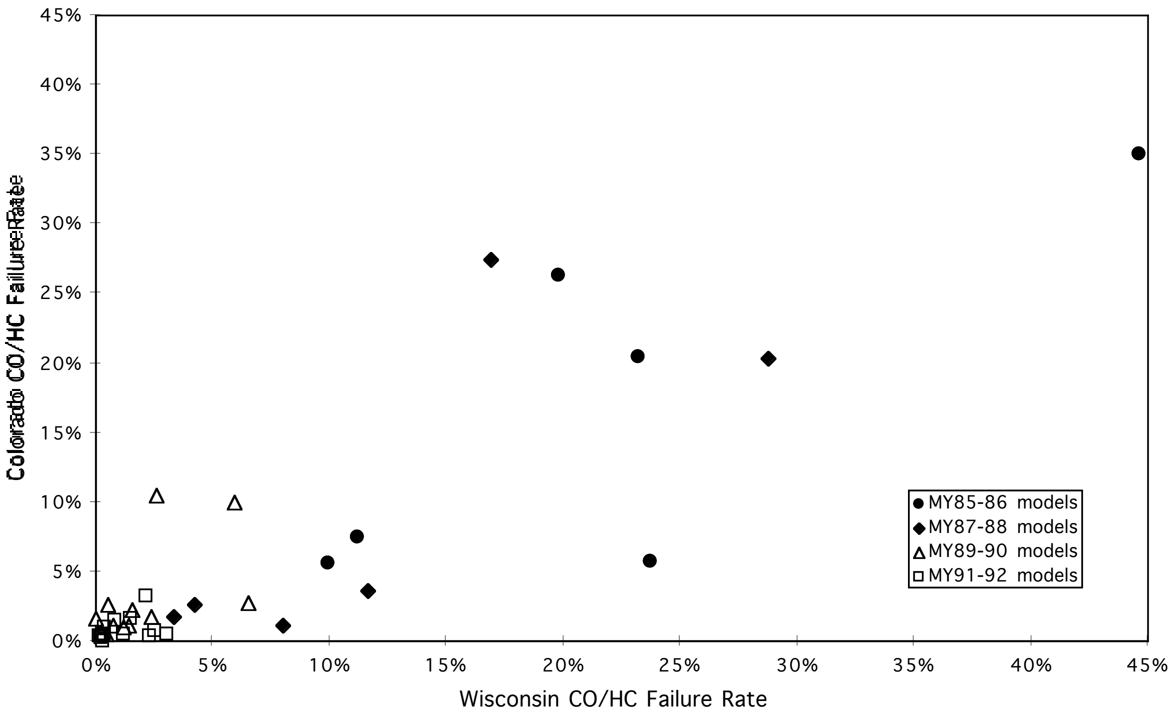


Figure 10. WI v. CO 1997 IM240 Combined CO and HC Failure Rates,
Selected Models



Figures 11 through 13 present a comparison of average emissions from 79 models from the Arizona and Ohio IM240 programs. Since Ohio only provided emissions test results, and not whether individual vehicles passed or failed the IM240, we compared the average emissions by model of vehicles receiving the full test. Unlike Arizona, Ohio does not allow dirty vehicles to fast fail; all failing vehicles are given the full IM240. As a result, there are many more vehicles tested per model in Ohio than in Arizona. Again, we limit the analysis to those models for which at least 10 individual vehicles received the full IM240. The different symbols in the figure distinguish models that have over 20 individual vehicles tested (diamonds) from models with 10 to 20 individual vehicles tested (open triangles).

The figures indicate good agreement in average emissions by model for two IM240 programs. For the most part, this agreement does not appear to weaken when we include the models with relatively few measurements of individual vehicles. The exception is a few models that have substantially higher average CO in the Arizona program than in the Ohio program (the four open triangles near the bottom right corner of Figure 12). Overall, average emissions by model are consistently lower in Ohio than in Arizona, even though: 1) the Ohio models are one to two years older at the time of testing (the Ohio data are more recent); 2) the majority of Ohio cars in our sample are failures, whereas at most about half of the Arizona cars are failures; and 3) Ohio had no previous I/M program prior to the IM240 program (Arizona had an idle program previously).

Summary

Further analysis of Arizona IM240 data indicates that the failure rates by vehicle model are quite consistent over multiple test years. An interesting finding is that the I/M program does not appear to reduce failure rates, particularly for CO; rather, the program merely offsets the expected rate of failure due to vehicle use. Analysis of failure rates by model, and by individual vehicle, over several years of IM240 testing is needed.

Adjusted average emissions by model from vehicles not receiving the full IM240 agree with average emissions by model from full IM240 tests. By developing a satisfactory method to project full IM240 emissions measurements from short test results, we can greatly increase the statistical power of the IM240 data. More work needs to be done to understand the biases in our simple adjustment methodology, and to improve the adjustment procedure.

The comparison of data from three IM240 states shows good agreement in failure rates by vehicle model. However, this analysis is based on a small number of models from several model years. A more detailed analysis of I/M failure rates from multiple states is needed. Preliminary comparison of the Arizona and Ohio IM240 programs indicates good agreement in average emissions by vehicle model. More study is needed to determine why emissions are consistently lower in Ohio.

Figure 11. Average HC by Model, AZ vs. OH
MY90-93 models with at least 10 full IM240 tests

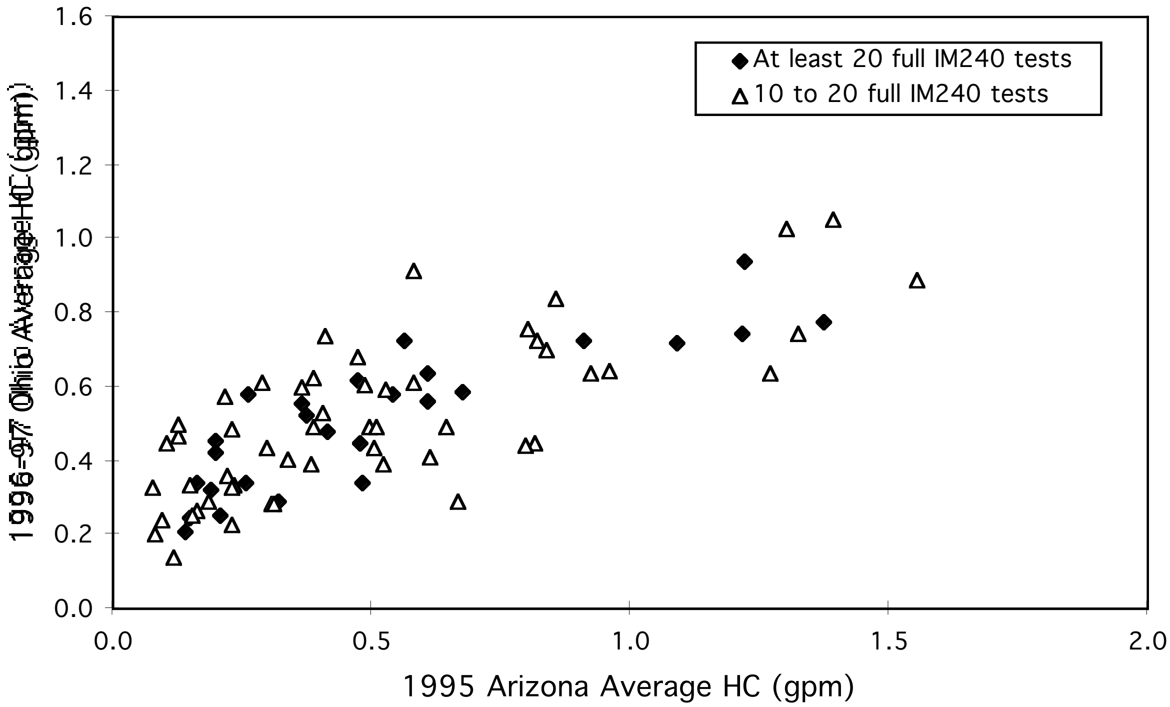


Figure 12. Average CO by Model, AZ vs. OH
MY90-93 models with at least 10 full IM240 tests

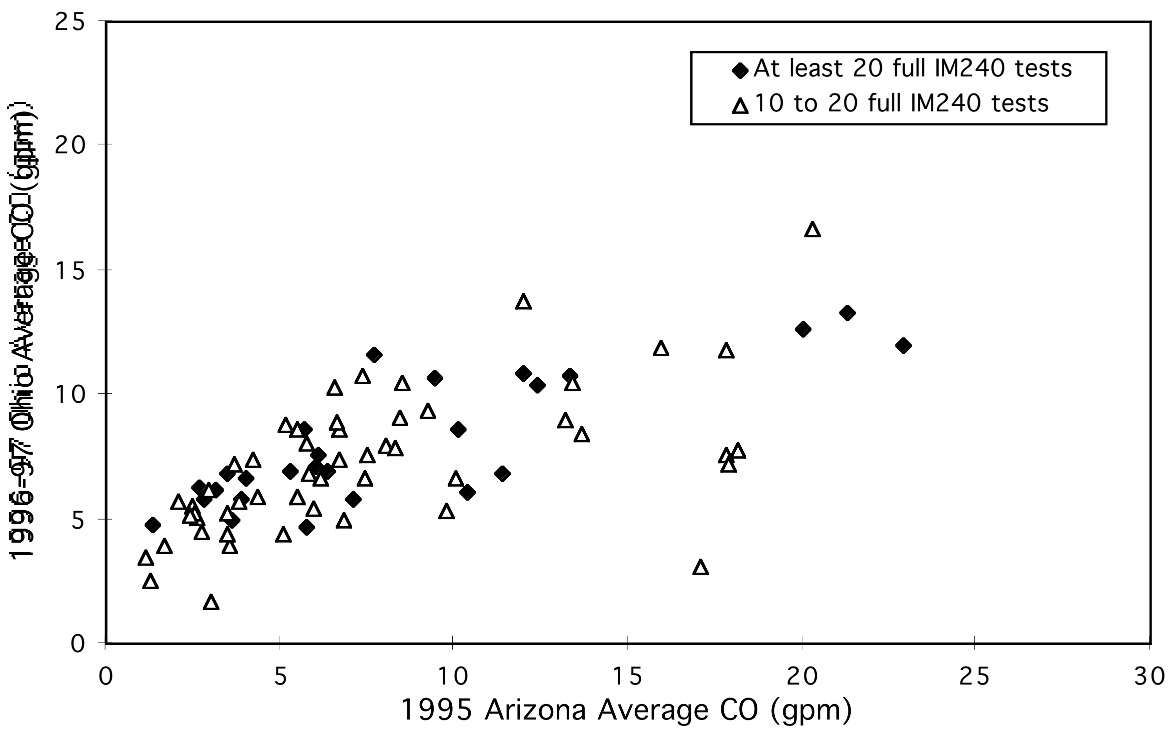
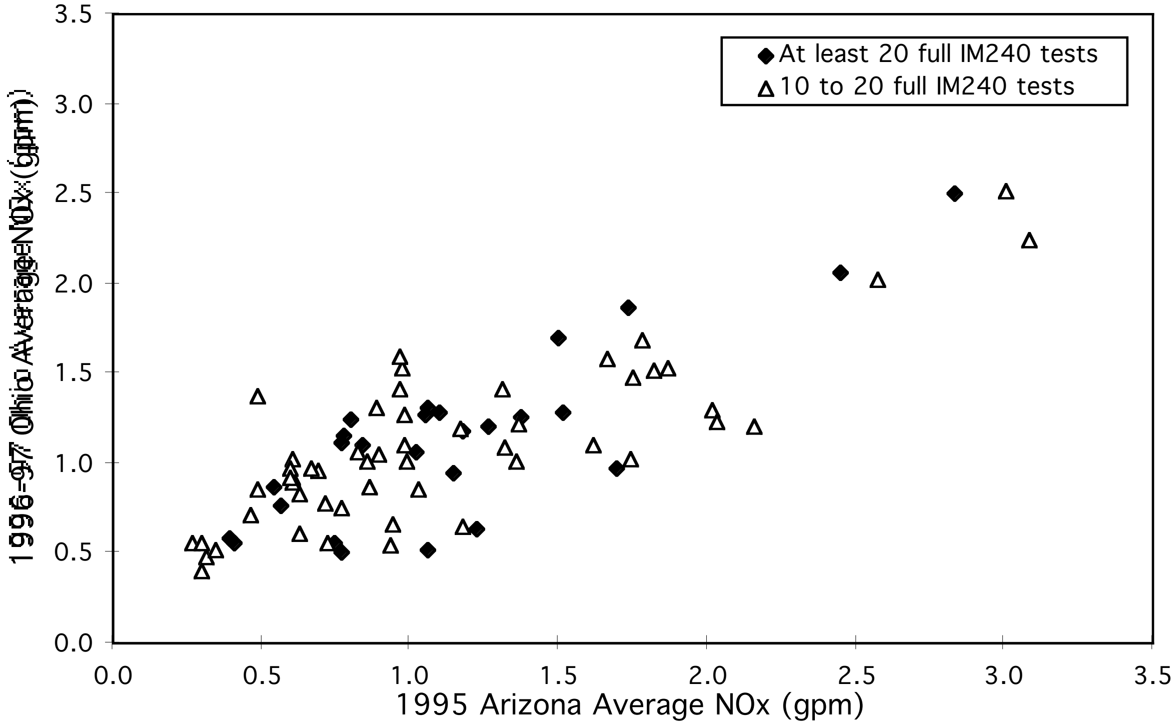


Figure 13. Average NOx by Model, AZ vs. OH

MY90-93 models with at least 10 full IM240 tests



APPENDIX G

In-Use Emissions by Vehicle Model

Tom Wenzel and Etan Gumerman, Lawrence Berkeley National Laboratory

Poster presented at the Ninth CRC On-Road Vehicle Emissions Workshop, April 21, 1999, San Diego CA.

Previous research indicates that there is a wide range in in-use emissions by vehicle model. Data on average emissions by vehicle model can be used for a variety of purposes, from identifying suspected low-emitting vehicles for exemption from I/M testing, to creating incentives for consumers to purchase, and manufacturers to build, vehicles with durable emissions controls. Last year we demonstrated the consistency in failure rate and average emissions by model year and model, using three years of data from the Arizona I/M program. We also presented a preliminary comparison of average emissions by vehicle model from several IM240 programs. This year we more thoroughly compare average emissions by vehicle model from the Arizona, Colorado, and Wisconsin enhanced I/M programs.

Elements of Three I/M Programs

Table 1 summarizes the key features of the enhanced I/M programs in Arizona, Colorado, and Wisconsin. Important differences are the cutpoints used (Arizona's and Wisconsin's are similar, while Colorado's tend to be less stringent), and the model years tested in each year (while Arizona tests all model years each year, Colorado tested mostly odd model years in 1997 and Wisconsin tested mostly even model years). Differences in the test cycles in Colorado and Wisconsin complicate analysis between the two programs. **Figure 1** demonstrates the difference in test cycles in the Colorado and Wisconsin programs. The figure shows the number of vehicles tested from July to December 1996 in all three states, by model year.

Similar numbers of vehicles from each model year were tested in Arizona in 1996, while the majority (90%) of vehicles tested in Wisconsin are from odd model years, and most (65%) of the vehicles tested in Colorado are from even model years. Colorado requires an I/M test when a vehicle is sold, and the next scheduled I/M test is not required until two years later. Therefore, most of the large number of vehicles from odd model years tested in 1996 were sold at some point earlier in their lifetime. (In contrast, vehicles sold in Wisconsin do not change their test cycle; the small number of even model year vehicles tested in 1996 in Wisconsin are early or voluntary tests.) In order to get large enough samples of vehicles from a particular model year in each state, we use 6 months of data from two calendar years, July 1996 to June 1997. **Figure 2** shows that this approach reduces the "sawtooth" effect due to different test cycles in Colorado and Wisconsin.

Figure 1. Number of Vehicles by Model Year and State

Passenger Cars, July 1996 to December 1996

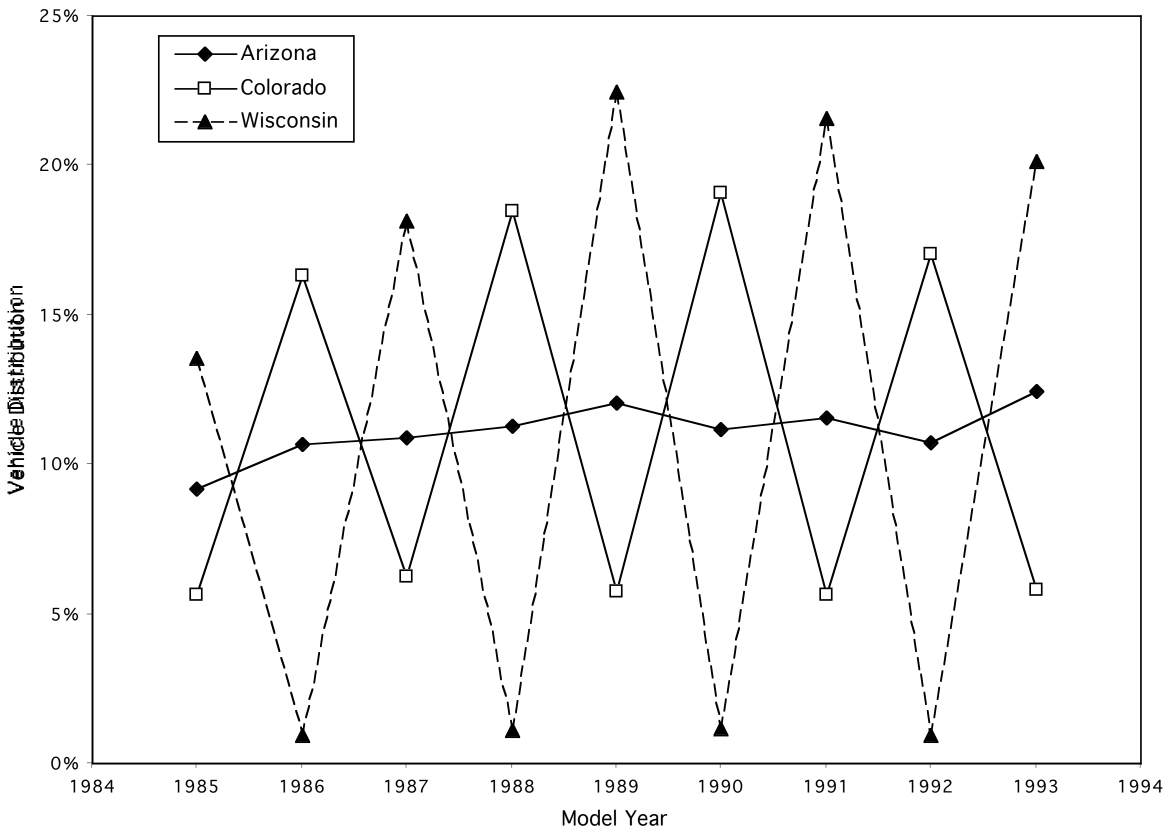


Figure 2. Number of Vehicles by Model Year and State

Passenger Cars, July 1996 to December 1996

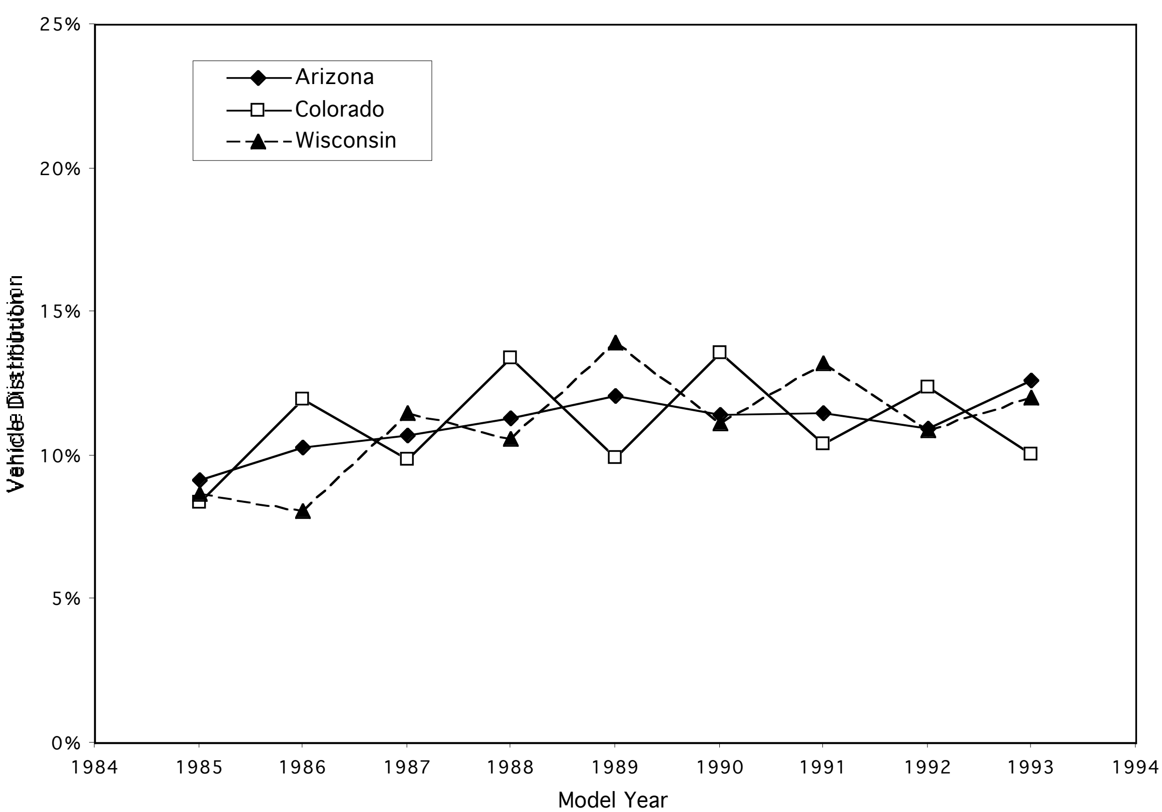


Table 1. IM240 Program Elements in Three States

Program Element	Arizona	Colorado	Wisconsin (1)
Test Cycle	biennial; all MYs tested in 1997	biennial; odd MYs tested in 1997	biennial; even MYs tested in 1997
Test on Resale?	no	yes	yes
Composite Cutpoints (cars)			
HC	91-95: 1.2 81-90: 2.0	86-95: 4.0 82-85: 5.0	91-95: 1.25 81-90: 2.0
CO	91-95: 20 83-90: 30 81-82: 60	91-95: 20 85-90: 25 83-84: 50 82: 65	91-95: 20 83-90: 30 81-82: 60
NOx	91-95: 2.5 81-90: 3.0	95: 4.0 86-94: 6.0 82-85: 8.0	*91-95: 2.5* *81-90: 3.0*
Fast Pass? Fast Fail? Phase 2 Pass?	yes yes yes	yes no no	yes no yes
Second Chance to Pass?	no	yes if emissions <2x cutpoint	yes if emissions <2x cutpoint
Full Tests	random 2%	all vehicles tested 1/97 to 3/97	1996: all vehicles tested on weekends; 1997: random 2%

(1) Cutpoints shown were effective 12/96 to 11/97. Although Wisconsin tests for NOx, vehicles are not failed for exceeding NOx cutpoints. Vehicles tested during weekends in 1996 were given full test; this practice was replaced by 2% random sampling in 1997.

(2) Adjusting Short Test Emissions to Full IM240 Equivalents

In our analysis we use average emissions rather than failure rate, since the emissions cutpoints differ among the states and many new car models have low failure rates. Within a state, average emissions by model correlate quite well with failure rate by model. A limitation of using average emissions is that IM240 testing procedures are not consistent between vehicles. All three states allow the cleanest vehicles to pass after 30 seconds of testing (fast passes); Arizona allows the dirtiest vehicles to fail after 94 seconds (fast fails), while Colorado and Wisconsin give all failing vehicles the full IM240 test.

To compare emissions from vehicles tested over different portions of the IM240, we need to correct fast-pass/fast-fail emissions to full test equivalent values. We use the same simple methodology to convert short test results in Arizona and Wisconsin to full test equivalent emissions. This methodology uses correction factors based on the average ratio of emissions at each second to full test emissions, for each pollutant and second of the IM240. Colorado uses a slightly different methodology to convert short test emissions to full IM240 equivalents; we use the Colorado adjustments for the vehicles tested in the Colorado program. For our purposes here, we do not require that this correction results in absolute accuracy for individual vehicles; rather we look for consistent ranking of models among the three states.

Arizona runs full IM240 tests on a random sample of two percent of the vehicles in the fleet; in Colorado, the fast-pass feature was "turned off" for all vehicles tested in the first three months of 1997 (that is, all vehicles tested during this period received a full IM240 test). We compare the average CO emissions of full tests with those of fast-pass/fast-fail tests, by vehicle model, in Arizona (Figure 3) and Colorado (Figure 4). The model year 1990 to 1993 car models shown have full tests on at least 10, and fast-pass/fast-fail tests on at least 250, individual vehicles.

Figure 3 indicates that there is no consistent bias in our adjustment procedure; average adjusted emissions by model from fast-pass/fast-fail tests in Arizona match very well with average emissions from full IM240s (perfect correlation between full tests and fast-pass/fast-fail tests is shown as a solid line, the actual correlation is shown as a dashed line). CO emissions from both fast-pass/fast-fail and full IM240 tests are higher in Colorado than in Arizona. As shown in **Figure 4**, the procedure to adjust Colorado fast-pass emissions appears to be somewhat biased. The Colorado procedure slightly overpredicts adjusted emissions from low emitting models, and slightly underpredicts adjusted emissions from the high emitting models. This is surprising, since the emissions from the highest-emitting vehicles, which have the biggest influence on average emissions of a particular model, are not adjusted, because all failing vehicles receive the full IM240 test in Colorado. Figures 3 and 4 indicate that the two different procedures used to adjust short test emissions to full IM240 equivalents give qualitatively similar results.

Figures 3 and 4 also show the value of using the average emissions values for vehicles receiving the short test. The vertical "whiskers" are the standard error associated with the full test cars, while the horizontal whiskers are the error of the fast-pass/fast-fail cars. The figures graphically demonstrate how an increase in the number of individual vehicles tested greatly reduces the statistical uncertainty of the average emissions of that model.

Figure 3. Average CO by Model, Fast Pass/Fail vs. Full Tests

Arizona MY90-93 with over 10 full IM240 tests (July 1996 -- June 1997)

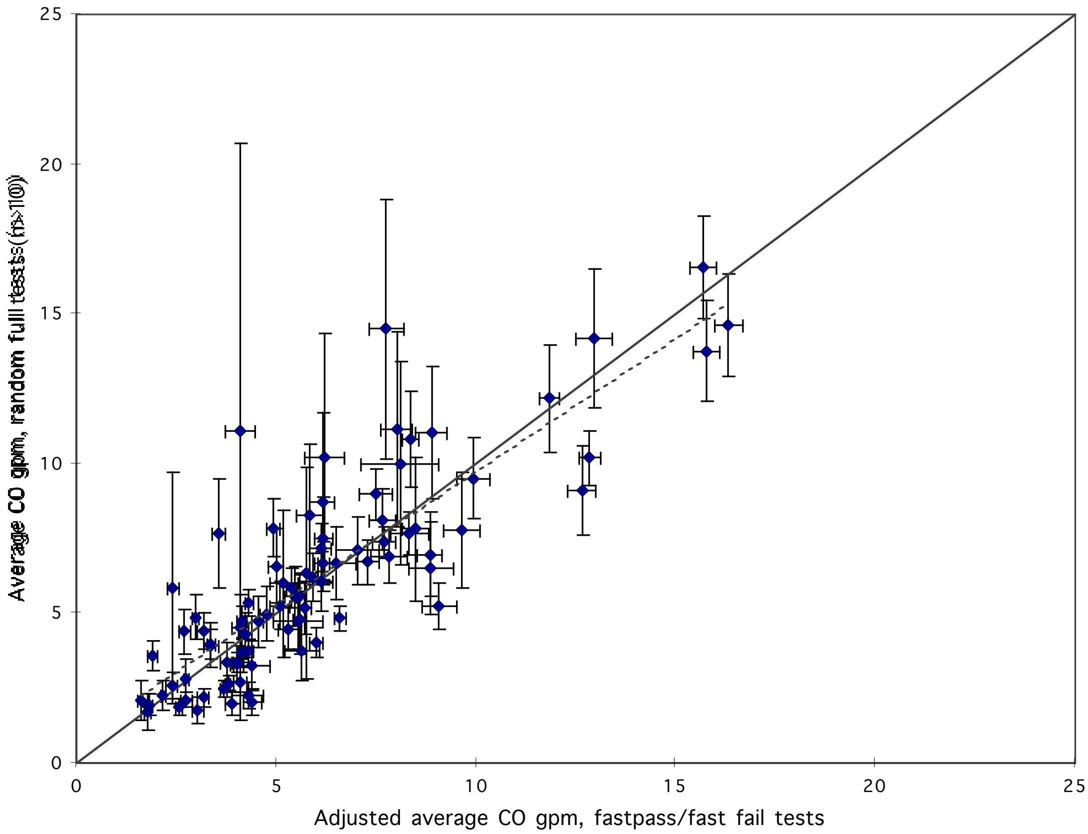
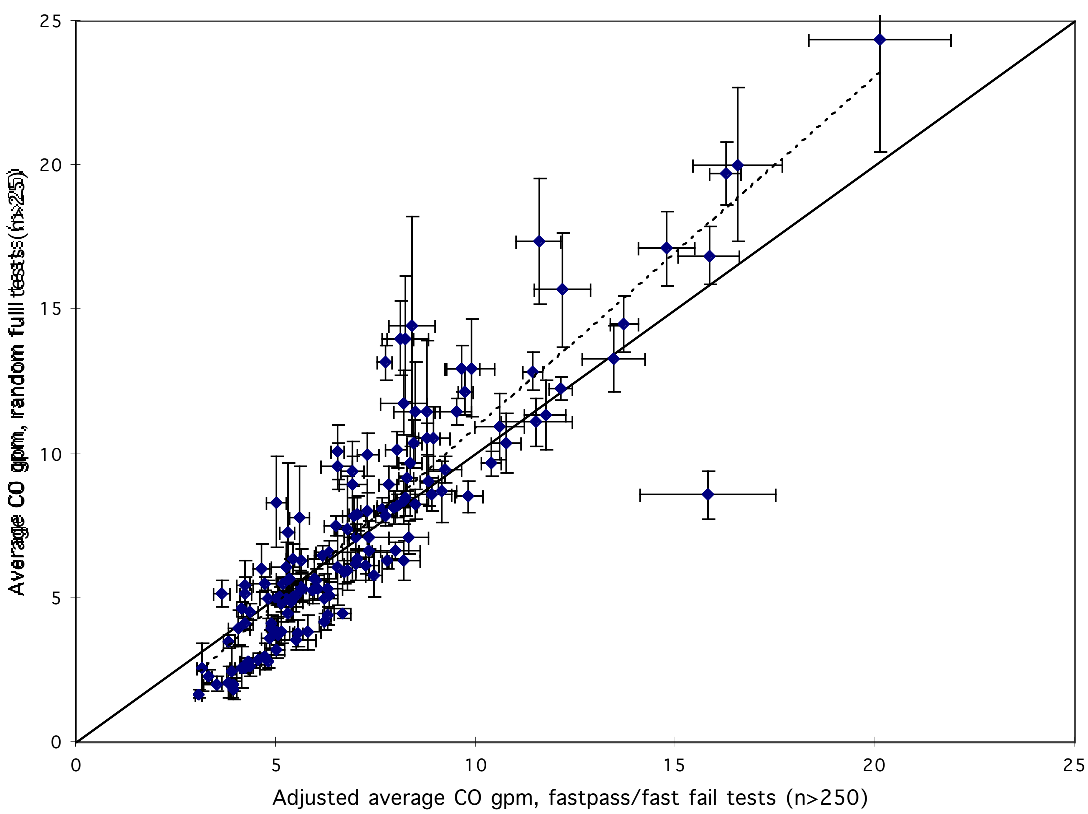


Figure 4. Average CO by Model, Fast Pass/Fail vs. Full Tests

Colorado MY90-93 with over 25 full IM240 test (July 1996 -- June 1997)



Average IM240 Emissions by Model in Three States

Figures 5 through 7 compare the average emissions of NOx, HC, and CO for 47 model year 1991 car models with at least 100 individual vehicles tested in each state. Each point represents a particular vehicle model, with average emissions from Arizona plotted on the x-axis and average emissions from Colorado and Wisconsin plotted on the y-axis. Average emissions by model in Colorado are designated by closed diamonds, whereas average emissions by model in Wisconsin are shown with open triangles. In each figure the solid line shows correlation with the Arizona data, while the dashed lines indicate the regression lines for the Colorado and Wisconsin data.

Figure 5 shows excellent agreement in average NOx by model among the three programs. NOx emissions are slightly higher in Arizona than in Colorado and Wisconsin. NOx emissions by model range from about 0.5 gpm to over 1.5 gpm, a factor of 3 difference between the lowest- and highest-emitting models. Two models are the highest emitters in each state, while 3 models are the lowest emitters in each state.

Figure 6 shows good agreement among the three states in terms of average HC by model. HC emissions are consistently lower in Wisconsin than in Arizona and Colorado. HC emissions by model range from about 0.2 gpm to over 0.8 gpm, a factor of 4 difference between the lowest- and highest-emitting models. Four models have consistently high emissions in all three states, while 6 models have consistently low emissions in all three states.

The two models with the highest emissions in Colorado, the Chrysler 2.2 liter and the Ford 5.0 liter, have relatively low emissions in Arizona and Wisconsin; these points are circled in the figure. One extreme emitter in Colorado, with 23 gpm HC, causes the average emissions for the Ford model to increase dramatically; removing this single vehicle reduces the average for that model to 0.77 gpm.

However, examination of the emissions distributions of these models also indicates that the difference in their average emissions among the states is due to generally higher emissions from many individual vehicles. **Figure 6a** compares the cumulative vehicle distributions for HC emissions from the Chrysler 2.2 liter model in the three states. The y-axis shows the cumulative fraction of vehicles with emissions above a given level on the x-axis; for example, about 8% of the vehicles in Colorado have HC emissions greater than 2.5 gpm, while less than 3% of the vehicles in Arizona have HC emissions greater than 2.5 gpm. The points noted indicate individual vehicles with high emissions. Even for the cleaner vehicles, the Chrysler 2.2 liter vehicles in Colorado have higher emissions than those in the other states; for example, 60% of the Colorado vehicles have HC greater than 0.5 gpm, while only 20% of the Wisconsin vehicles have HC above 0.5 gpm. Also, the dirtiest 1% of vehicles in Colorado (4 cars) have HC emissions nearly twice that of the dirtiest 1% of vehicles in Wisconsin (8 cars) and Arizona (4 cars).

Figure 6b compares the Colorado HC emission vehicle distributions of the two outlier models with those of a model that has consistently high HC emissions in each state (Saturn SL/SC MFI) and a model that has consistently average emissions in each state (Nissan Sentra). The figure illustrates that in a rank comparison vehicle by vehicle, every Nissan car has lower emissions than every Saturn. Consequently, the high average emissions of the Saturn model are a result of consistently high emissions across all Saturns, rather than a few individual vehicles with very high emissions. The two outlier models may exist either as a result of sensitivities in these particular models to differences in the state I/M programs, or due to other differences between the states that affect emissions. For example, perhaps the emissions controls of these models are

Figure 5. Average NOx by Car Model in Three States

MY91 Passenger Cars with at least 100 IM240 tests (July 1996 -- June 1997)

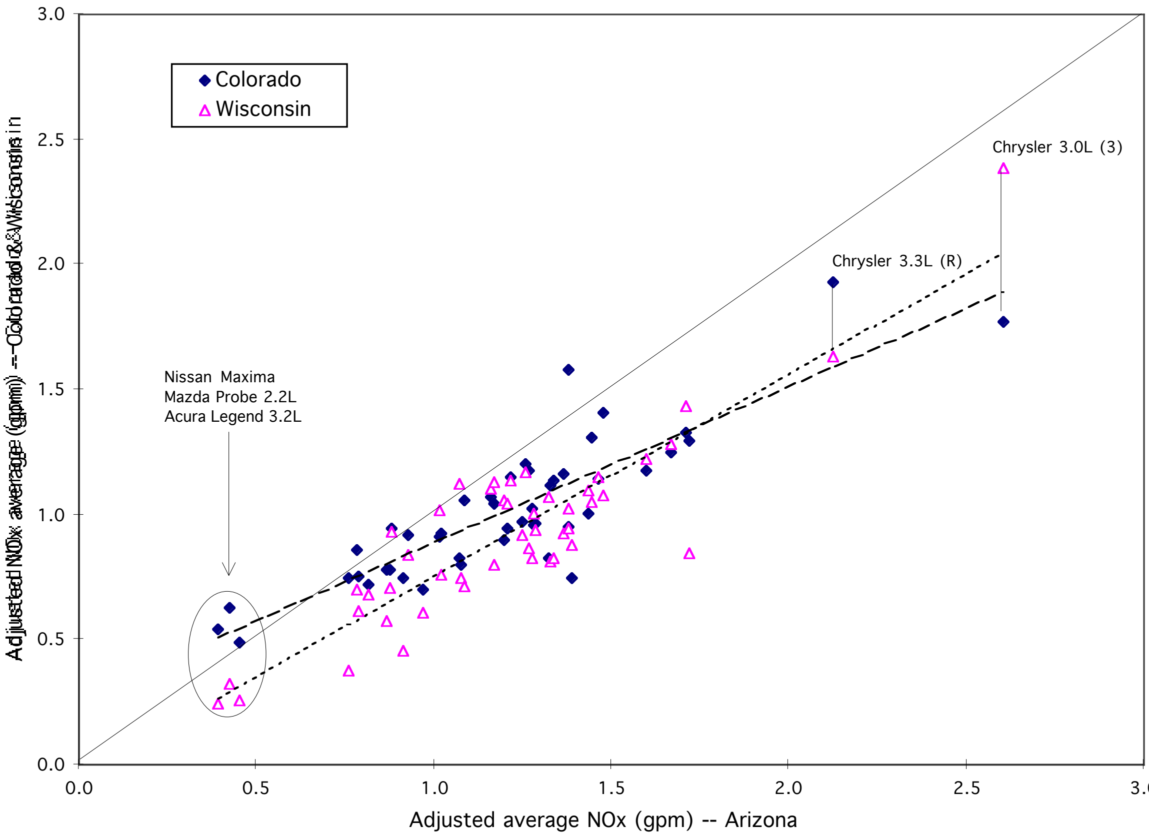


Figure 6. Average HC by Car Model in Three States

MY91 passenger Cars with at least 100 IM240 tests (July 1996 -- June 1997)

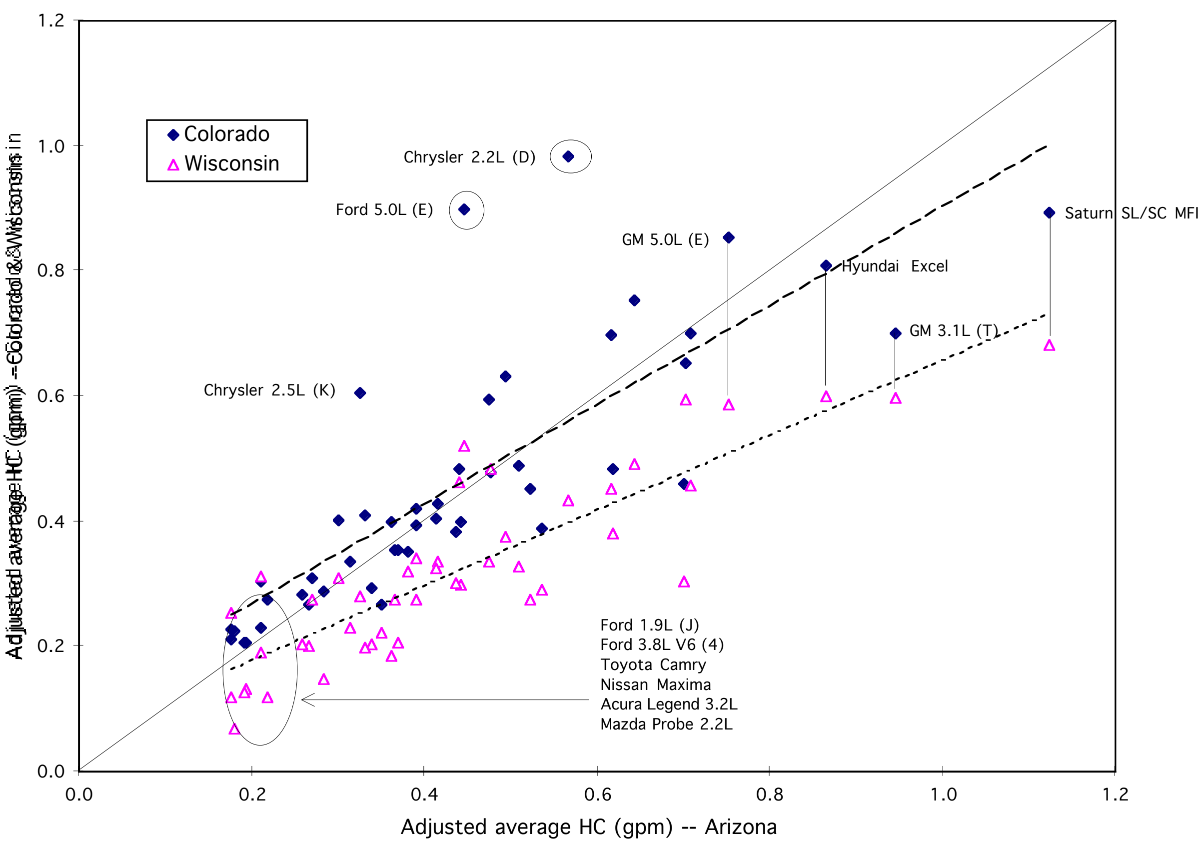


Figure 6a. Cumulative Vehicle Distribution for HC by State

MY 1991 Chrysler 2.2L (D) (July 1996 --June 1997)

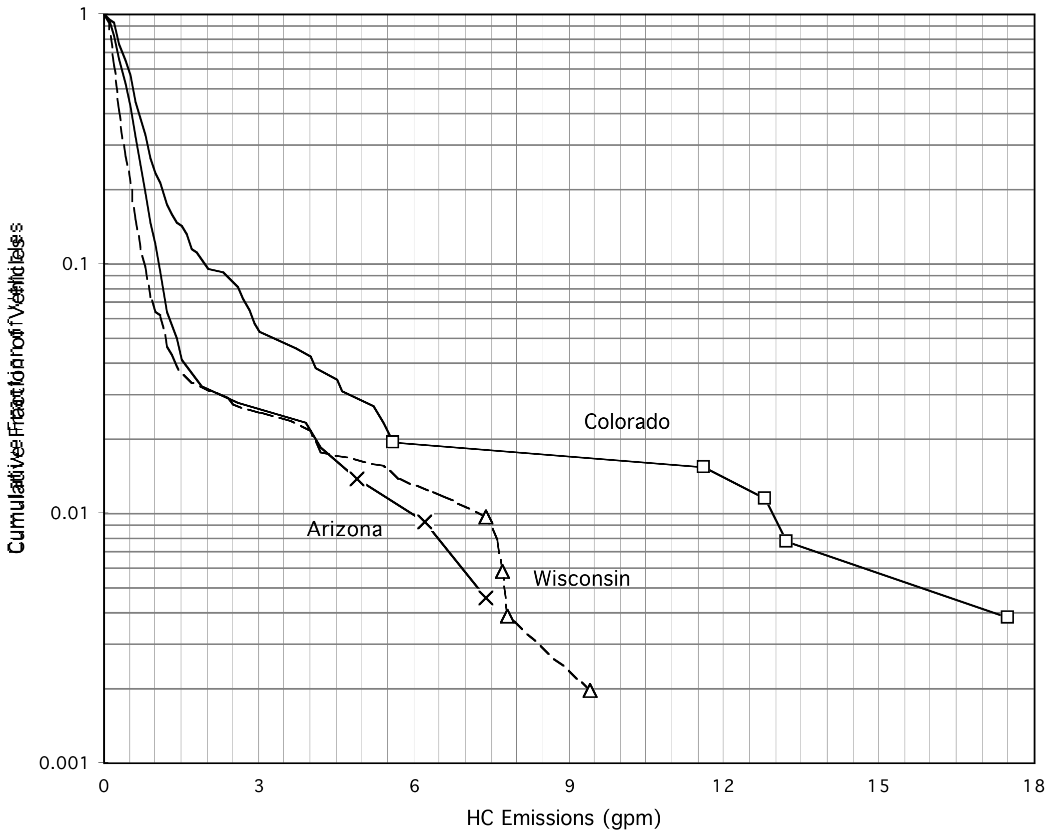
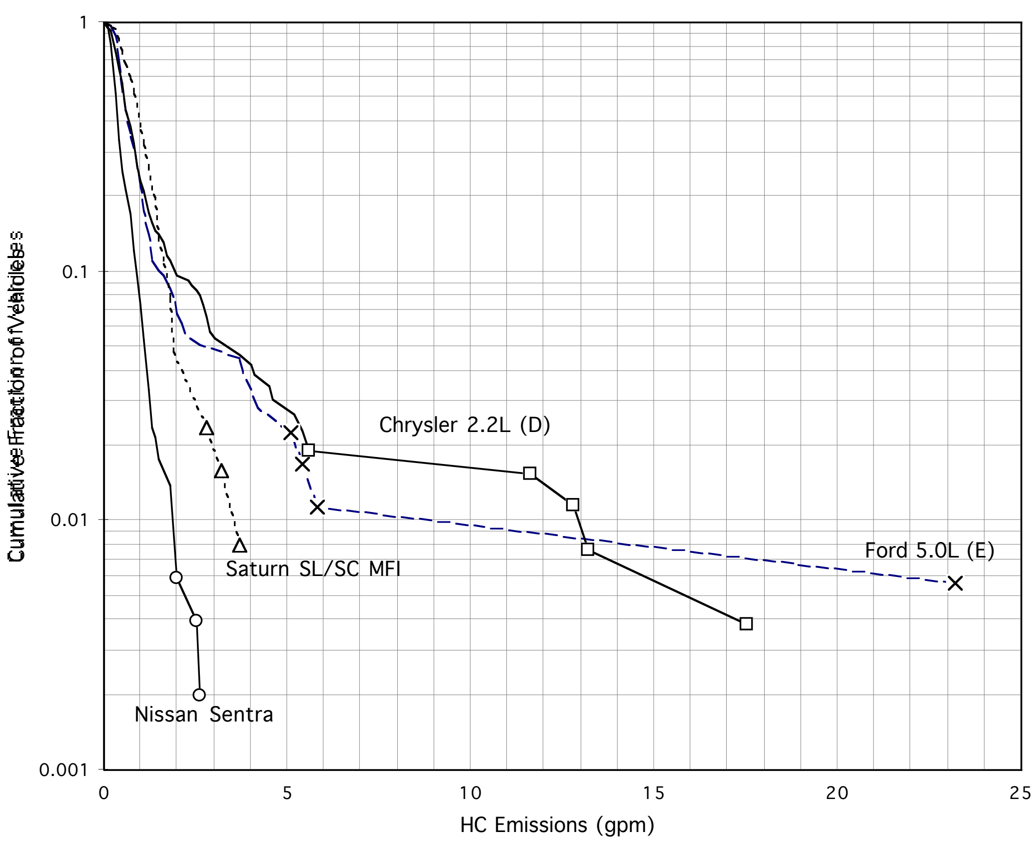


Figure 6b. Cumulative Vehicle Distribution for HC in Colorado

4 MY 1991 Passenger Cars (July 1996 -- June 1997)



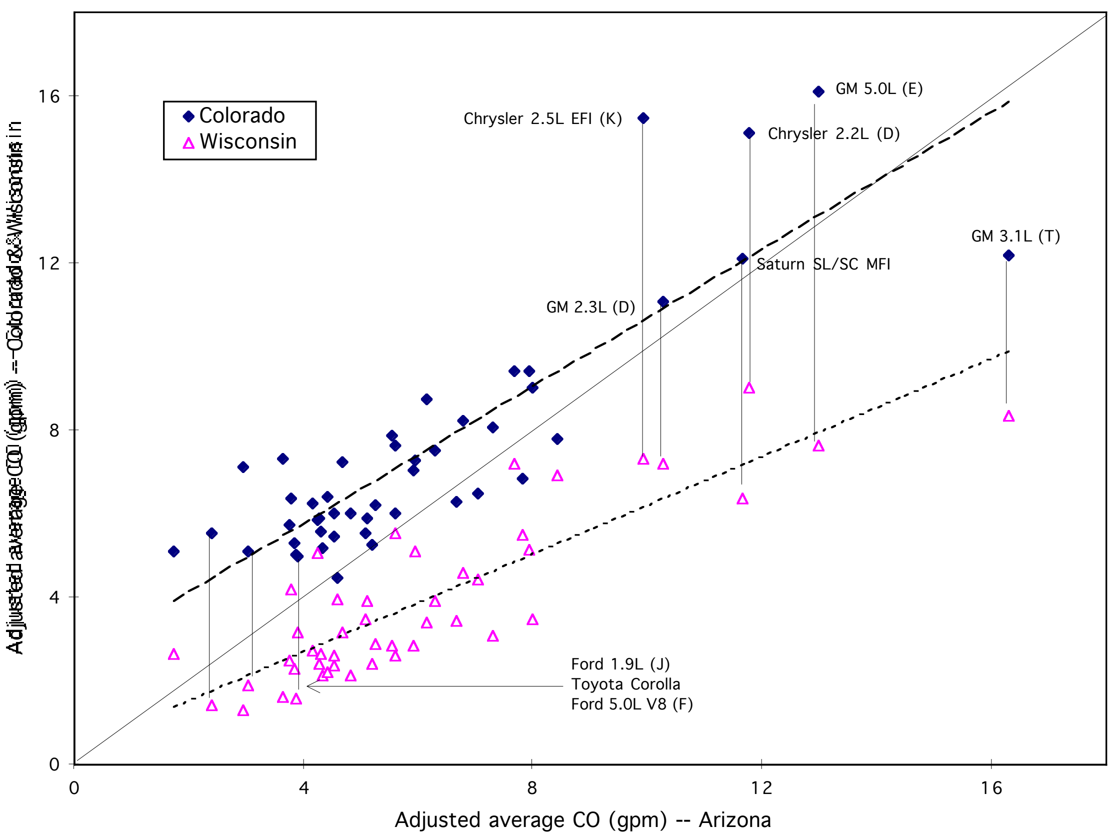
more sensitive to high altitude and/or high load driving, and therefore have higher emissions in Colorado than in Arizona or Wisconsin.

Figure 7 shows that average CO emissions for any given model tend to be substantially lower in Wisconsin, and substantially higher in Colorado, than in Arizona. Even so, there is good agreement among the three states. CO emissions by model vary by a factor of 3 in Colorado, to a factor of 7 in Wisconsin. Six models have consistently high emissions in all three states, while 3 models have consistently low emissions in all three states.

One possible explanation of the high Colorado, and low Wisconsin, CO emissions may be the different test cycles used in each state. Virtually all of the 1991 models were tested in 1996 in Wisconsin, while most of these models were tested in 1997 in Colorado; therefore, the Colorado vehicles are 6 months older on average than the Wisconsin vehicles. To evaluate this potential bias, we compared average emissions by model from vehicles tested between June 1996 and December 1996 only, and found that the Colorado CO emissions were reduced only slightly. There are two other factors that could account for the consistently higher emissions in Colorado: other differences in the I/M testing conditions, practices, or cutpoints used in each state, or differences in driving patterns, maintenance practices, and/or fuel composition in the three states that result in actual differences in in-use emissions.

Figure 7. Average CO by Model in Three States

MY91 Passenger Cars with at least 100 IM240 tests (July 1996 -- June 1997)



Summary

A comparison of in-use emissions data from three state IM240 programs indicates that average emissions by vehicle model are quite consistent across state programs. Several models are consistently among the cleanest, and the dirtiest, in each of the three states. Although the agreement is best for NOx, the comparisons for HC and CO are quite good. The two models with the highest HC emissions in Colorado have relatively low HC emissions in Arizona and Wisconsin. The inconsistent results for these particular models may be due to their sensitivity to I/M program differences, or to other factors that can affect in-use emissions.

APPENDIX H

Converting Fast Pass/Fast Fail Emissions Results to Full IM240 Equivalents
Tom Wenzel, Lawrence Berkeley National Laboratory
June 17, 1999

At least 4 methods have been used to convert fast-pass/fail emissions to full IM240 emissions: 1) a method developed by LBNL for use in analyzing the Arizona I/M program (LBNL; Wenzel, 1997); 2) a method developed by Peter McClintock of Applied Analysis, with input from Rob Klausmeier and the Colorado Department of Public Health and Environment, for use in the Colorado I/M program (PM; McClintock, 1998)[1]; 3) a method developed by Resources for the Future, also for use in analyzing the Arizona program (RFF; Ando et al, 1998), and 4) a method developed by EPA using data from Wisconsin and applied to Ohio fast pass data EPA; (EPA; Enns, 1999, and personal communication). The LBNL method is based on the average ratio of emissions at each second to full test emissions from a sample of 4,000 vehicles receiving the full IM240 in Arizona in 1992.[2] The LBNL method involves dividing emissions at a given second by a correction factor, based only on the second of testing (and not on other variables, such as vehicle age or type). The McClintock and RFF methods are similar; they rely on regression models generated for many subsets of the data. The McClintock method accounts for vehicle age and type, while the RFF method accounts for vehicle age and the product of the vehicle age and emissions level at a given second. The McClintock coefficients were calculated for 10 second intervals in the IM240 trace; coefficients for interlying seconds are determined by interpolation. The RFF method estimates <u>negative</u> emissions values for some vehicles with very low emissions at second 31. The EPA method is based on a single regression equation for the entire vehicle fleet. The equation includes coefficients for the log of fast pass emissions, the last second of the test, and dummy variables for whether the vehicle is fuel injected or carbureted, a car or a truck, and for 14 model years.[3]

This memo examines the accuracy of such methods in predicting full IM240 emissions. First, we compare the accuracy of each of the three methods on a sample of vehicles whose full test emissions are known. Then we apply the PM and LBNL method to Wisconsin data, to see what effect different methods have on fleet emissions estimates. Finally, we evaluate the accuracy of the LBNL method by comparing the distribution of emissions of fast pass/fast fail vehicles with that of the random sample of vehicles receiving the full IM240.

Comparison of Three Methods

We used the random sample (Jan-June 1996) of vehicles given a full IM240 in Arizona to test the accuracy of three different methods in accurately predicting full IM240 emissions from vehicles passing after only 30 seconds of testing. We first identified which vehicles in the random sample would have passed EPA-recommended fast-pass cutpoints at second 30; there are 2,197 such passenger cars in the random sample. Then, we calculated what each vehicle's

1. Developed in late 1995 and early 1996 with inputs from Rob Klausmeier and CDPHE.
2. The testing was conducted by Automotive Testing Laboratory, under contract with EPA. The vehicles tested may not be a random sample of vehicles.
3. We could not perfectly match EPA's results when we applied EPA's methodology to the Ohio data. We calculated the Ohio fleet emissions to be 18% higher for HC, 7% higher for CO, and 8% higher for NOx than as calculated by EPA, apparently using the same conversion method.

estimated full IM240 emissions would be under each conversion method. We analyzed cars from model years 1983 to 1990, and model years 1991 and newer, separately, since different fast pass cutpoints are applied to these two model year groups. (A more thorough analysis would predict at which second each vehicle would have fast-passed or fast-failed the IM240, and then make the adjustments to all of the vehicles in the sample. We focus here on the vehicles that fast-pass at second 30 to simplify the analysis, and because the majority of vehicles that fast-pass pass at this second.)

Table 1 shows the average emissions for these groups of vehicles over the full IM240 test, as measured under the program and as estimated by the three conversion methodologies. (The EPA method is calculated for MY81-94 cars only; the analysis was not applied to the 297 MY95 and newer cars in the Arizona sample. The EPA method was not applied to 77 cars in the MY83-90 group for which type of fuel delivery system was not readily available for the Arizona data. Restricting the analysis to only those cars that can be analyzed using the EPA method does not change the results.) The method that best predicts the emissions for each vehicle group and pollutant is noted in bold type in the table. In general, the LBNL method tends to underestimate the full test emissions of fast-passed cars; this underestimation is greatest for CO emissions, and for emissions from older vehicles. On the other hand, the PM method tends to overestimate emissions. The RFF method predicts emissions from older cars more accurately than from newer cars, while the LBNL method predicts emissions from newer cars more accurately. The RFF method estimates that full test NOx emissions of MY91+ cars are only 30% of their measured emissions; this large underestimation is because the RFF method predicts that over 35% of these vehicles would have had negative NOx emissions over the full IM240 (rounding the emissions of these vehicles to zero raises the fleet NOx emissions to 0.14 gpm, and raises the ratio of estimated to measured NOx to 0.31). When applied to Arizona data, the EPA method drastically underestimates emissions of all three pollutants, in both model year groups.

Table 1. Measured and Predicted full IM240 Emissions under Four Prediction Methods

	Average emissions, gpm			**Ratio of estimated to measured**		
	HC	CO	NOx	HC	CO	NOx
MY83-90 (n=1,204)						
Measured	0.42	6.85	1.21	1.00	1.00	1.00
LBL	0.30	3.65	0.87	0.72	0.53	0.72
PM	0.53	9.55	1.14	1.27	1.40	**0.95**
RFF	0.46	7.05	0.91	**1.11**	**1.03**	0.76
EPA*	0.12	2.19	0.33	0.29	0.32	0.27
MY91+ (n=993)						
Measured	0.10	1.93	0.45	1.00	1.00	1.00
LBL	0.10	1.35	0.43	**1.01**	0.70	0.95
PM	0.15	3.35	0.46	1.45	1.74	**1.01**
RFF	0.08	1.56	0.11	0.79	**0.81**	0.24
EPA**	0.05	0.76	0.23	0.49	0.40	0.51

*The EPA method relies on type of fuel delivery system (carbureted or fuel injected); the analysis was not applied to 77 cars for which fuel delivery system was not readily available.

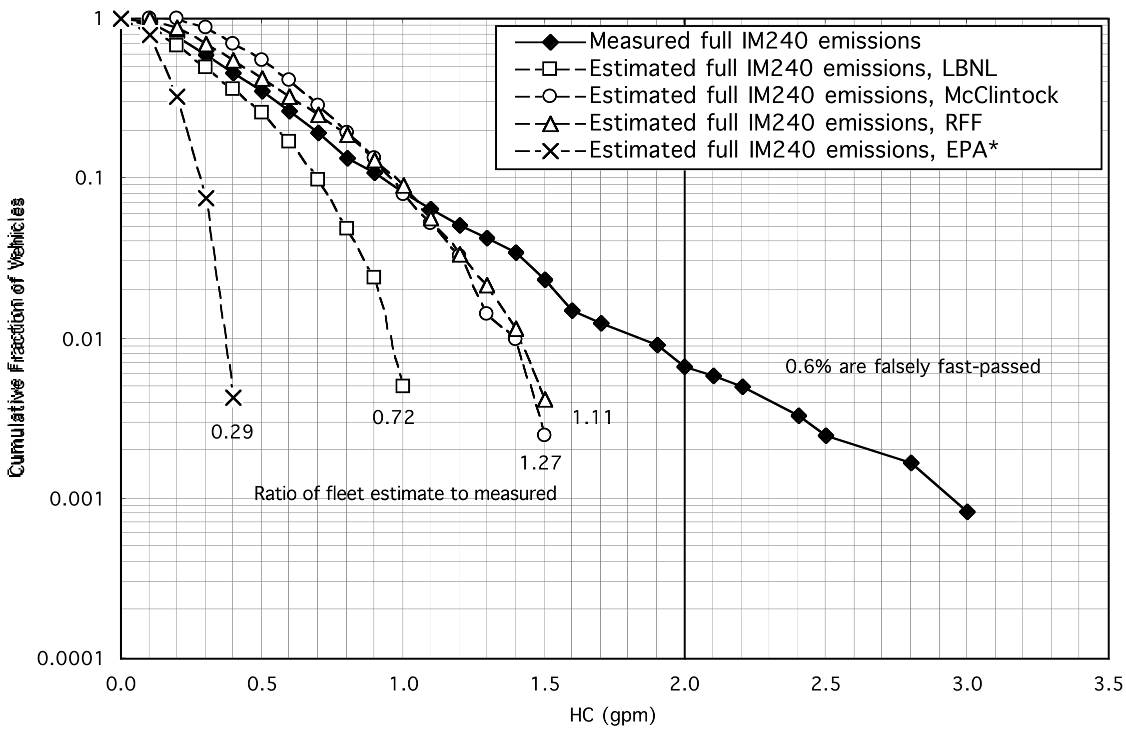
**The EPA method is calculated for MY81-94 cars only; the analysis was not applied to the 297 MY95 and newer cars in the Arizona sample.

Figures 1 through 6 show the distribution of emissions, as measured and as estimated based on the three prediction methods. The figures also report the ratio of the estimated to the measured emissions for all vehicles, from Table 1 above. A few of the vehicles that would have been fast-passed (i.e. that had emissions at second 30 lower than the fast pass cutpoints) had emissions higher than the cutpoints applied to the full IM240 test. The figures indicate the portion of all vehicles that would have been falsely fast-passed if the fast-pass cutpoints were applied. For instance, none of the MY91 and newer cars would have been fast-passed for NOx (Figure 6), but 2 percent (24 cars) of the MY83-90 cars would have been falsely fast-passed for NOx (Figure 3).

It should be noted that the RFF method was developed using some of the data used in this evaluation, and therefore should be expected to most accurately predict full test emissions. (The PM method was developed using Colorado IM240 data, the LBNL method was developed using earlier IM240 tests conducted in Tucson by Automotive Testing Laboratories, and the EPA method was developed using Wisconsin IM240 data.)

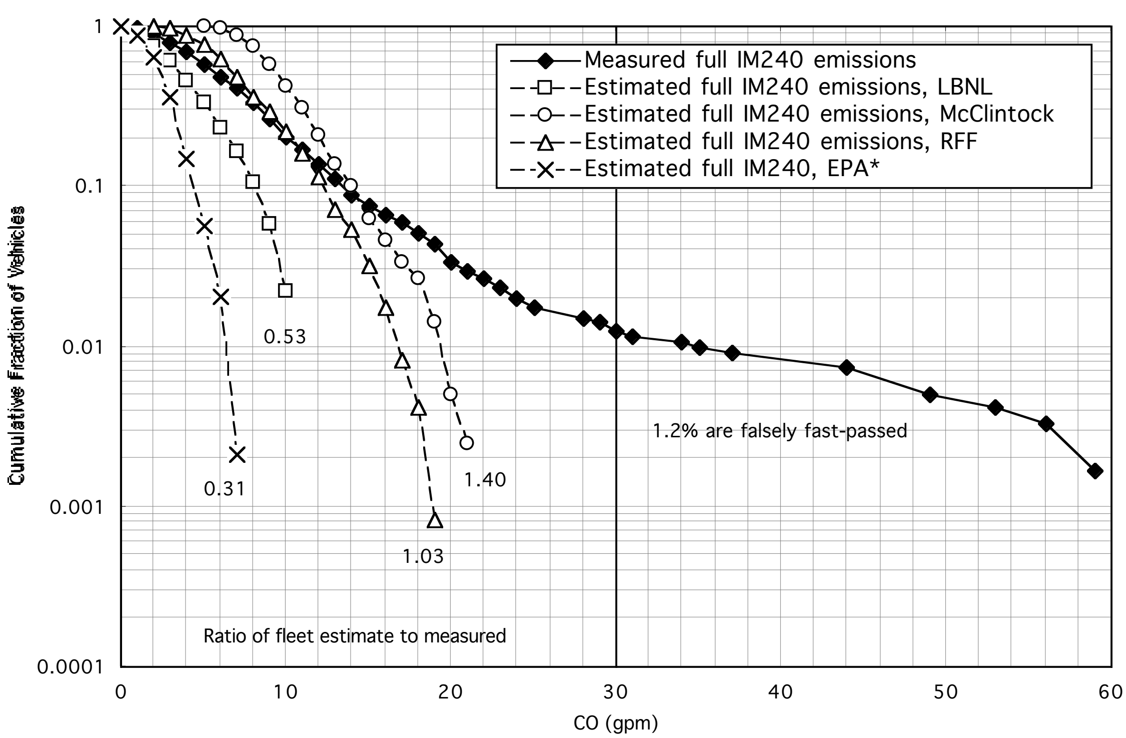
HC Distribution for Fast-Passed Vehicles

1204 MY83-90 cars passing start-up FP cutpoint, 1/96-6/96 Arizona IM240



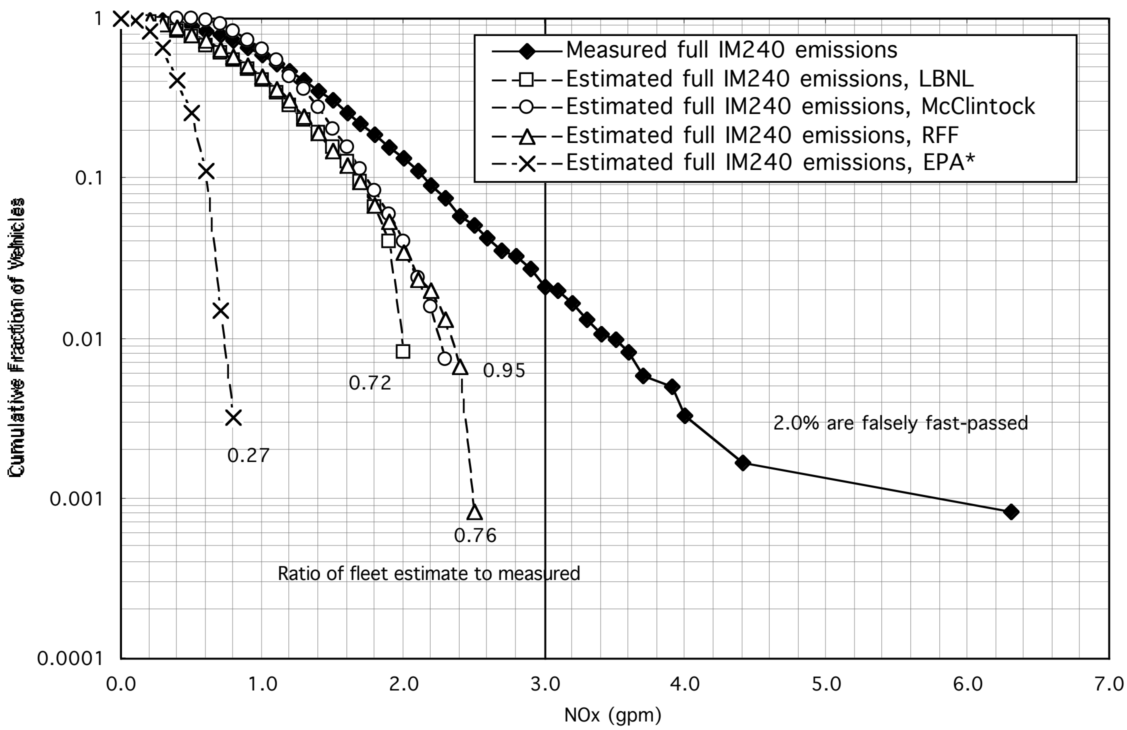
CO Distribution for Fast-Passed Vehicles

1204 MY83-90 cars passing start-up FP cutpoint, 1/96-6/96 Arizona IM240



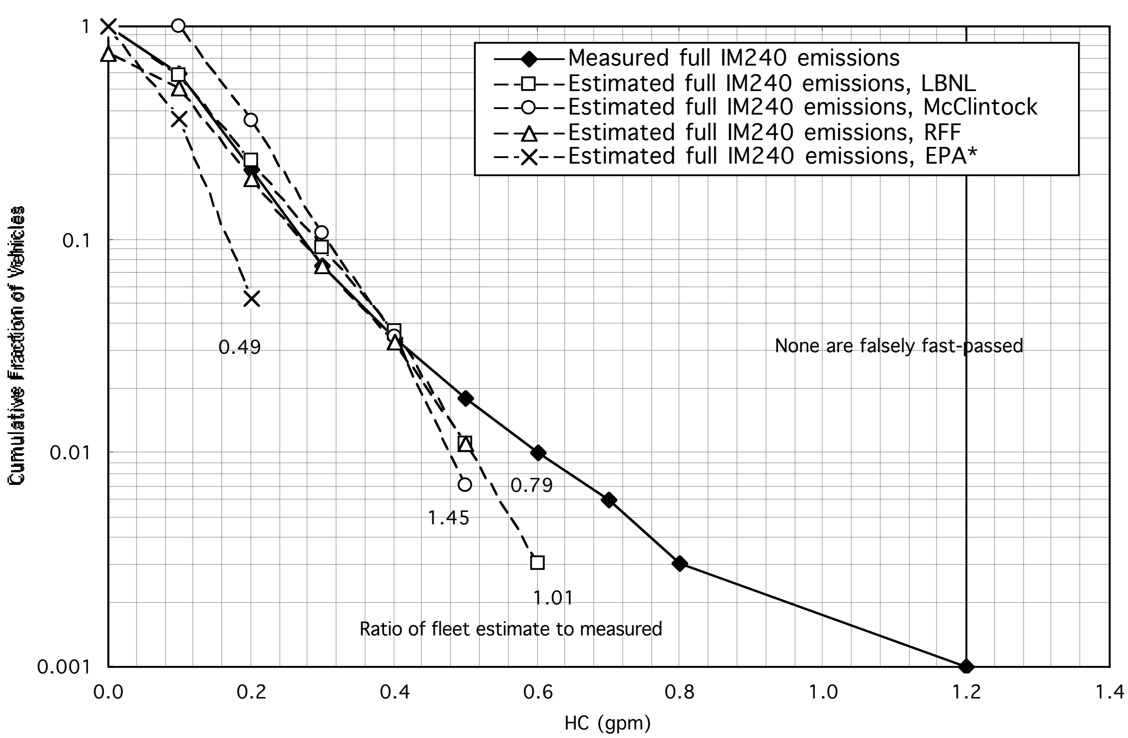
NOx Distribution for Fast-Passed Vehicles

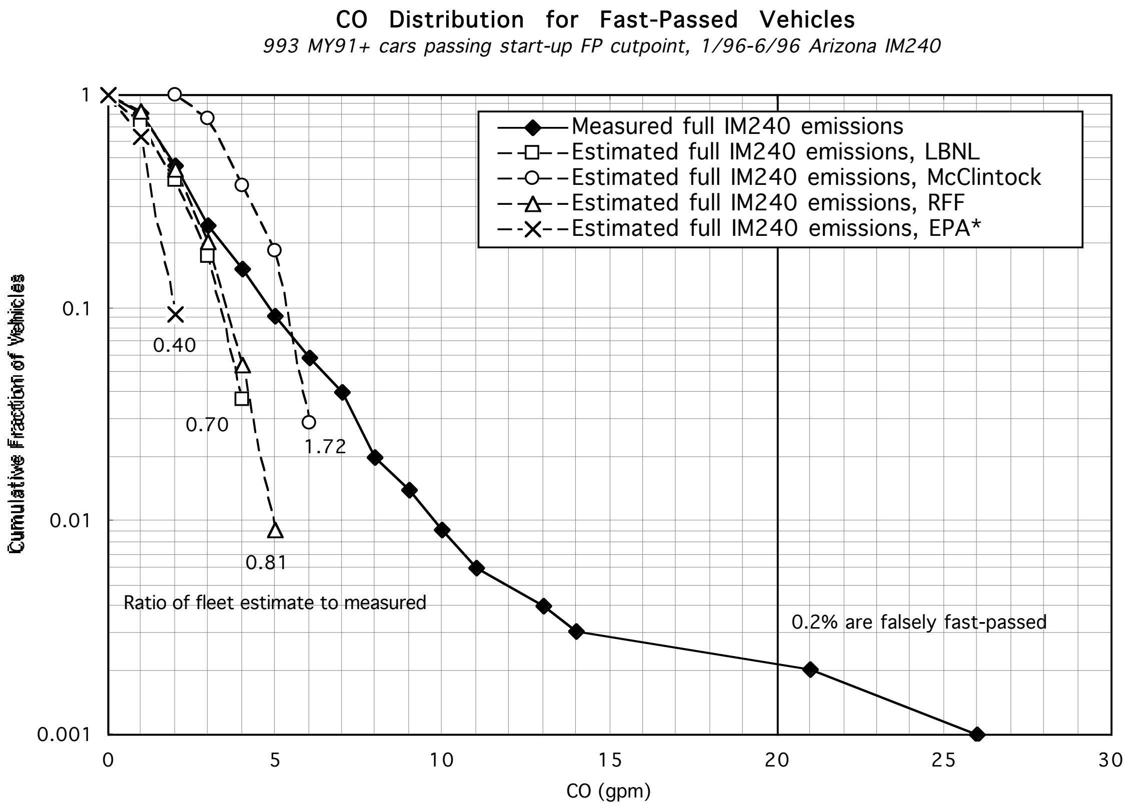
1204 MY83-90 cars passing start-up FP cutpoint, 1/96-6/96 Arizona IM240

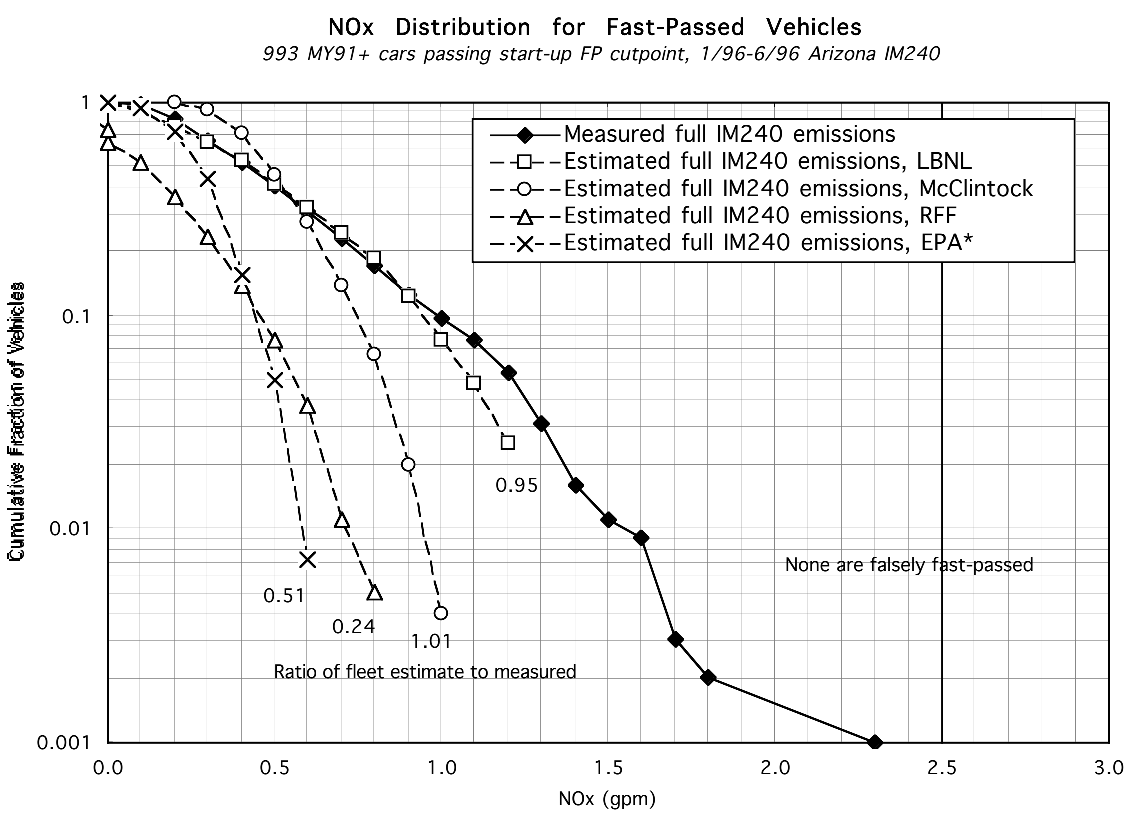


HC Distribution for Fast-Passed Vehicles

993 MY91+ cars passing start-up FP cutpoint, 1/96-6/96 Arizona IM240




CO Distribution for Fast-Passed Vehicles
993 MY91+ cars passing start-up FP cutpoint, 1/96-6/96 Arizona IM240
Measured full IM240 emissions
Estimated full IM240 emissions, LBNL
Estimated full IM240 emissions, McClintock
Estimated full IM240 emissions, RFF
Estimated full IM240 emissions, EPA*
Cumulative Fraction of Vehicles
CO (gpm)
0.40
0.70
1.72
0.81
Ratio of fleet estimate to measured
0.2% are falsely fast-passed


NOx Distribution for Fast-Passed Vehicles
993 MY91+ cars passing start-up FP cutpoint, 1/96-6/96 Arizona IM240
Measured full IM240 emissions
Estimated full IM240 emissions, LBNL
Estimated full IM240 emissions, McClintock
Estimated full IM240 emissions, RFF
Estimated full IM240 emissions, EPA*
Cumulative Fraction of Vehicles
NOx (gpm)
0.95
None are falsely fast-passed
0.51
0.24
1.01
Ratio of fleet estimate to measured

Two Methods Applied to Wisconsin Data

In order to determine the effect of using a different adjustment methodology on fleet average emissions, we applied the LBNL, PM, and EPA methods to an independent set of data from the Wisconsin IM240 program.[4] We also used the PM method based on a random sample of full tests conducted in Wisconsin, using data supplied by Peter McClintock. Table 2 shows the average emissions for the MY82 to MY94 passenger car fleet predicted by each method, as well as the ratio of the prediction under each method to the prediction under the PM method derived from Wisconsin data. The source of the data used for each method is listed in parentheses in Table 2. We only applied the data to vehicles for which we could identify their type of fuel delivery system, as the EPA method relies on this information. By restricting the analysis to these vehicles, we ensure that each method is applied to the same vehicles.

The LBNL method consistently predicts lower fleet emissions than the PM (Wisconsin) method, particularly for cars passed after only 30 seconds of testing. On the other hand, the EPA method predicts slightly higher fleet emissions than the PM (Wisconsin) method, especially for HC and NOx. The PM method based on Colorado data predicts the same fleet HC emissions as the PM (Wisconsin) method, but predicts higher CO emissions and lower NOx emissions. This type of analysis only tells us the relative effect of each prediction method on fleet emissions; we cannot say which method is more accurately predicting full test emissions.

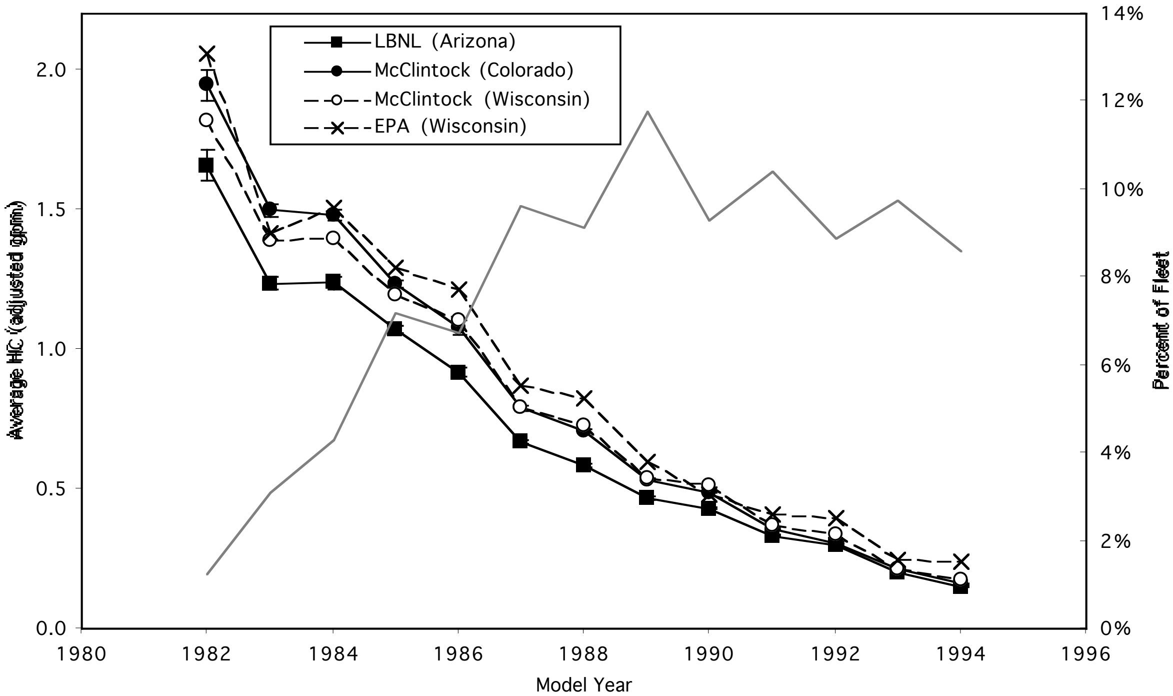
Table 2. Comparison of Different Methods on Wisconsin Data

	Average Predicted Emissions, gpm			**Ratio of Prediction to PM (Wisconsin) Prediction**		
	HC	CO	NOx	HC	CO	NOx
All tests						
LBNL (Arizona)	0.56	6.64	1.08	0.87	0.79	0.82
PM (Colorado)	0.64	9.94	1.15	**1.00**	1.18	0.87
PM (Wisconsin)	0.64	8.45	1.32	1.00	1.00	1.00
EPA (Wisconsin)	0.70	9.13	1.46	1.10	**1.08**	**1.11**
Cars passed after only 30 seconds of testing						
LBNL (Arizona)	0.23	2.26	0.71	0.62	0.45	0.68
PM (Colorado)	0.37	6.84	0.86	**1.00**	1.36	0.82
PM (Wisconsin)	0.37	5.02	1.04	1.00	1.00	1.00
EPA (Wisconsin)	0.43	5.22	1.18	1.15	**1.04**	**1.13**

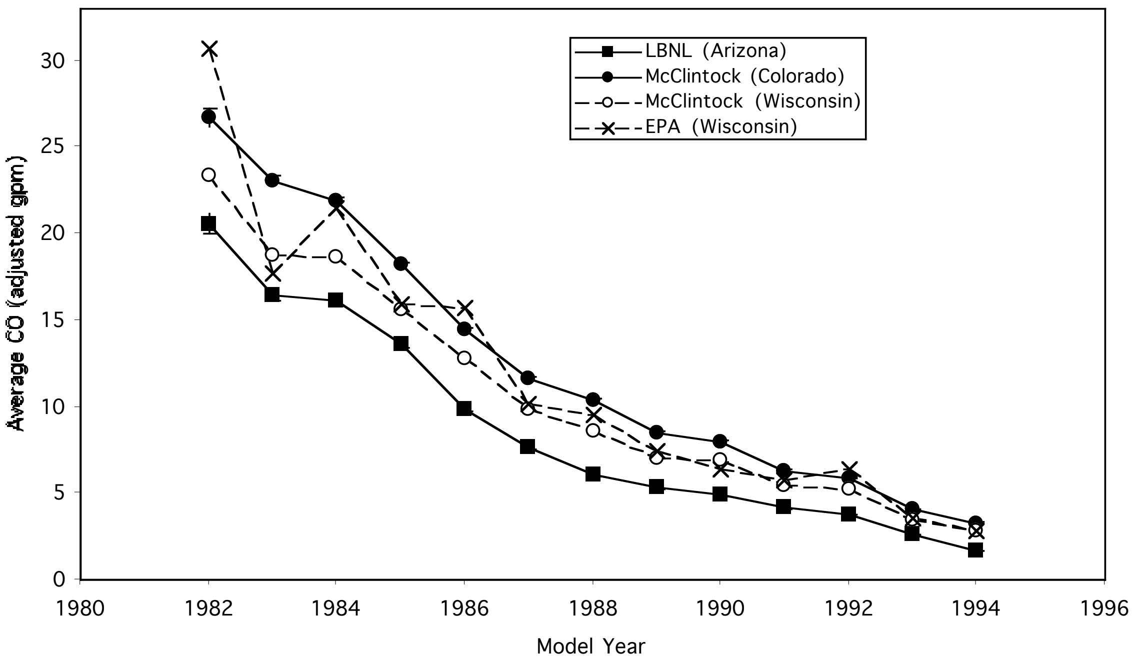
Figures 7 through 9 compare the average adjusted emissions for all MY82 to MY94 passenger cars by model year, under each prediction method. Figure 7 shows the percent distribution of cars by model year, as a gray line. Both of the methods based on Wisconsin data (the PM and EPA methods) result in higher emissions from even-year vehicles; this is particularly evident for NOx under the EPA method. These peaks are likely due to the sample of vehicles given the full test in Wisconsin, which were used to develop the adjustment methods. Most of this testing was conducted in 1996; therefore, most of these vehicles were from odd model years. McClintock's method results in smaller peaks because he grouped several model years together before calculating his adjustment factors.

4. Like Colorado, Wisconsin's IM240 program does not allow vehicles to fast fail; all vehicles with high emissions are given a full IM240.

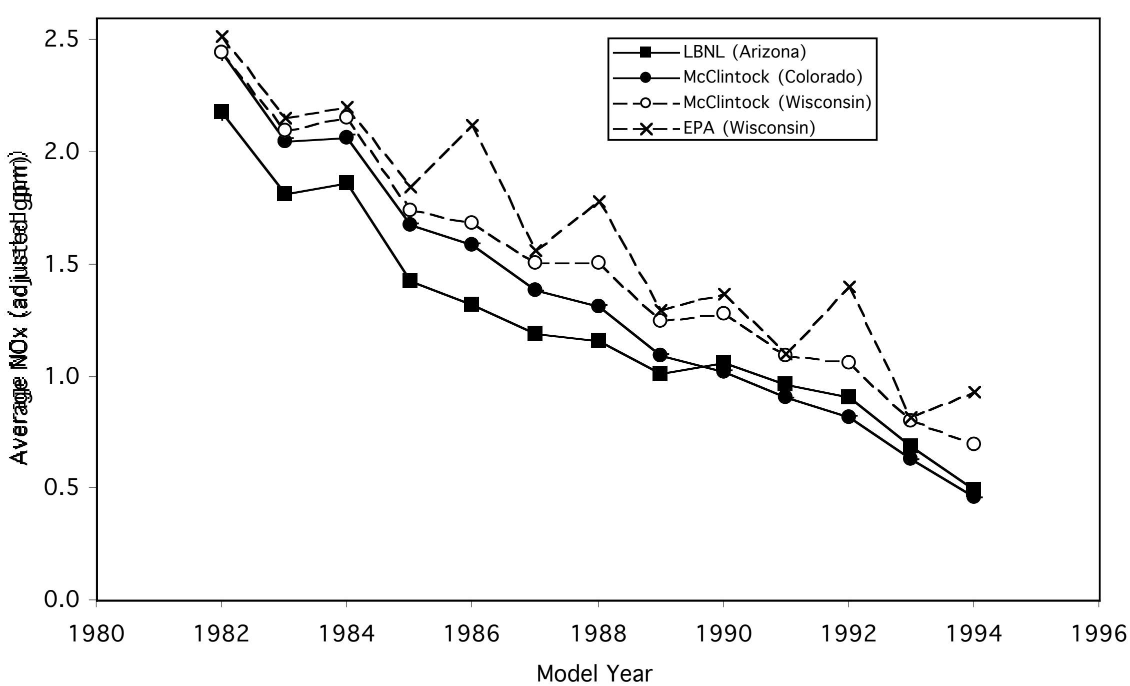
Average HC by Fast Pass Correction Factor and MY
MY82-94 Passenger Cars, Wisconsin 1996-97 IM240s



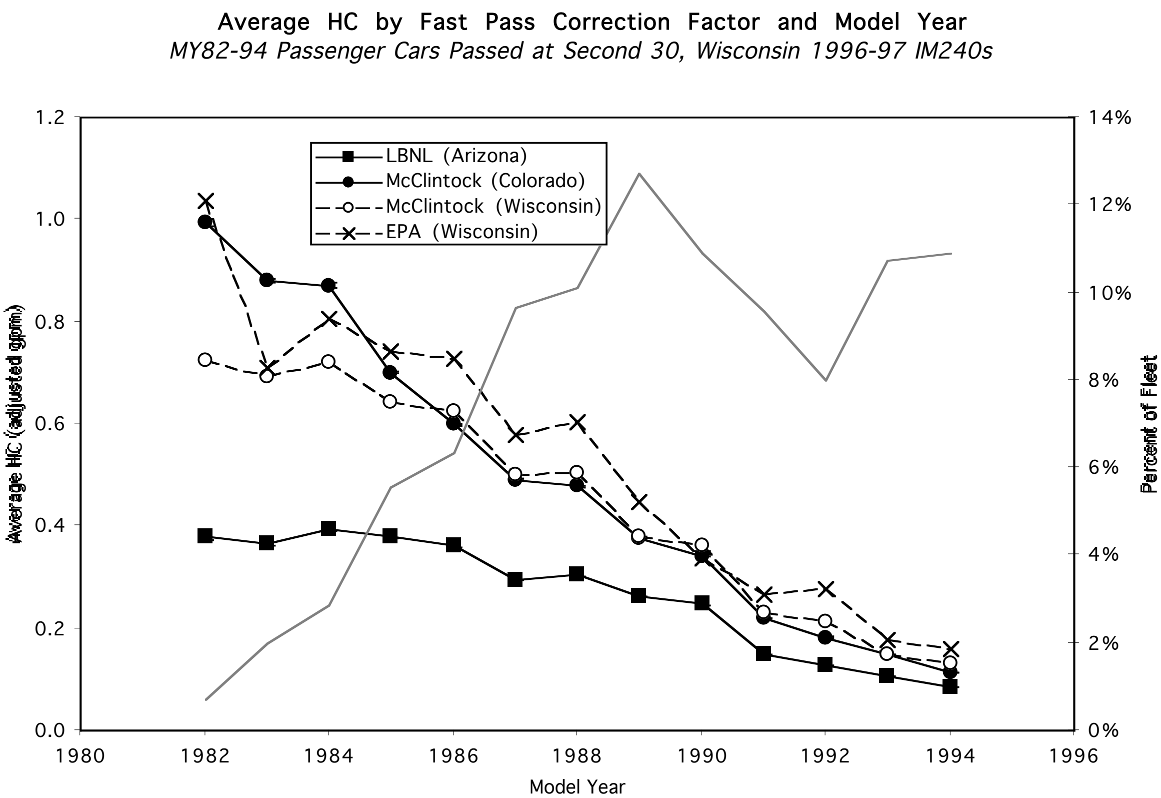
Average CO by Fast Pass Correction Factor and MY
MY82-94 Passenger Cars, Wisconsin 1996-97 IM240s



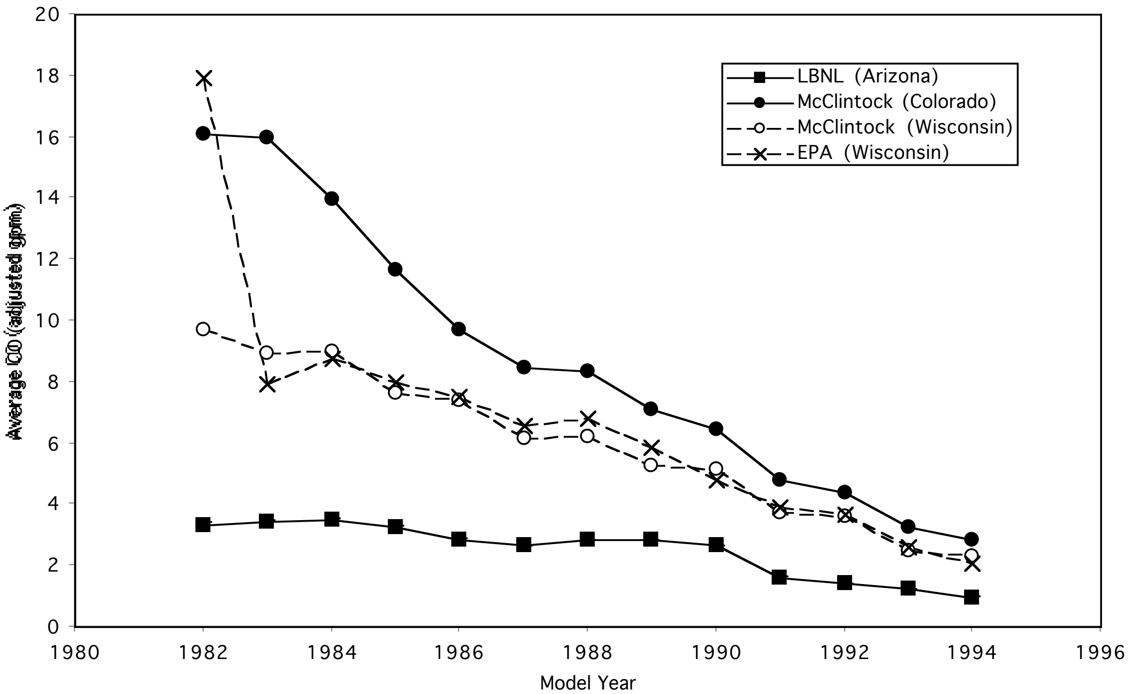
Average NOx by Fast Pass Correction Factor and MY
MY82-94 Passenger Cars, Wisconsin 1996-97 IM240s



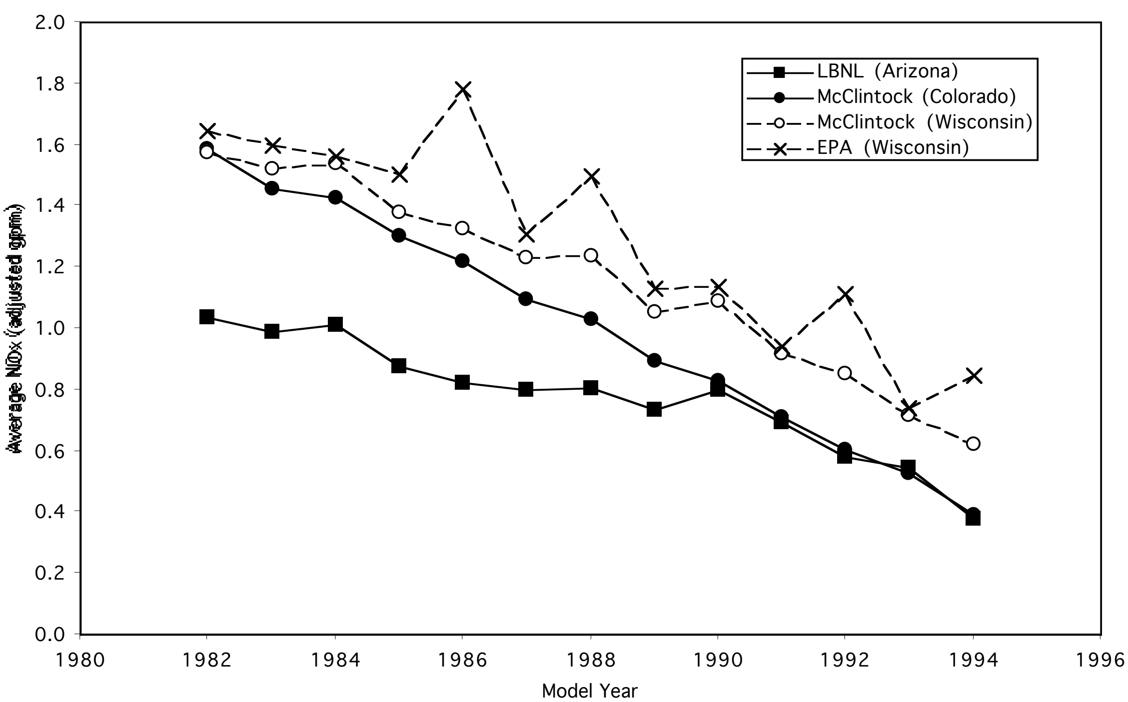
Figures 10 through 12 compare the average emissions by model year for vehicles that are fast-passed after only 30 seconds of testing. The accuracy of an adjustment method after only 30 seconds of testing greatly affects the overall accuracy of the method, since most vehicles are passed at this time. In this sample nearly 70% of all cars were passed after only 30 seconds. Here we see much larger discrepancies between the LBNL method and the PM (Wisconsin) method, especially for older cars. Again, Figure 10 shows the percent distribution of cars by model year, as a gray line.



Average CO by Fast Pass Correction Factor and Model Year
MY82-94 Passenger Cars Passed at Second 30, Wisconsin 1996-97 IM240s

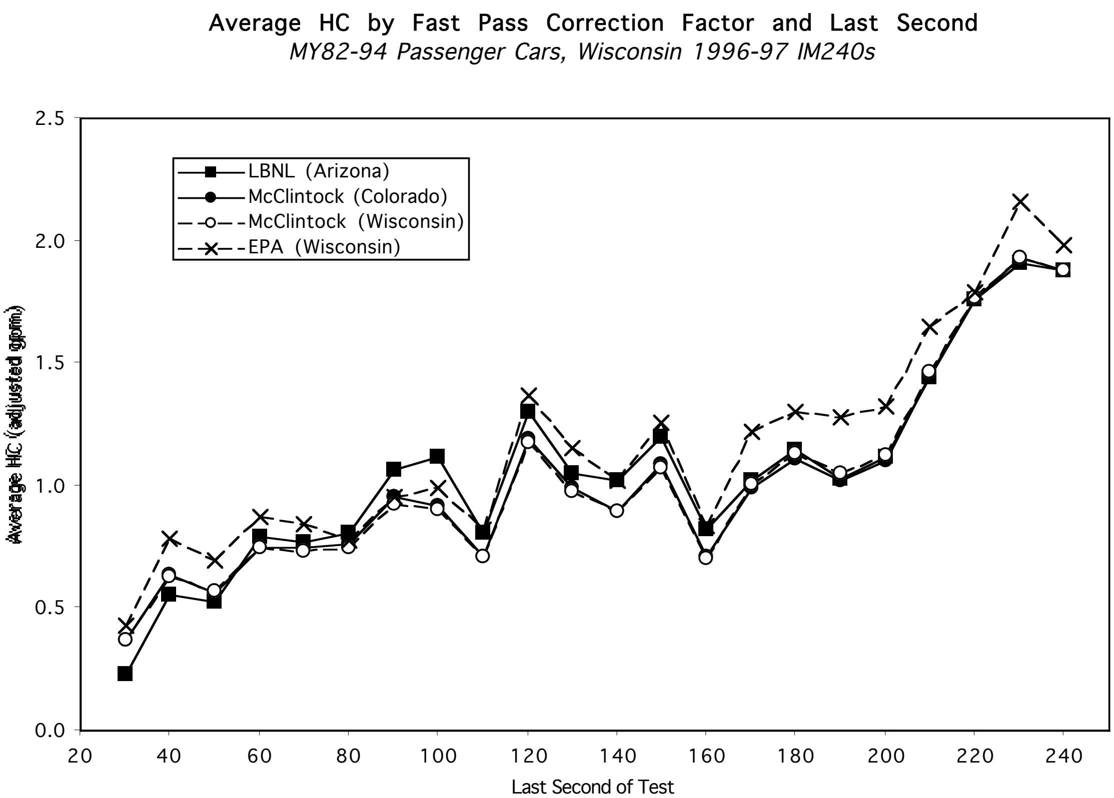


Average NOx by Fast Pass Correction Factor and Model Year
MY82-94 Passenger Cars Passed at Second 30, Wisconsin 1996-97 IM240s



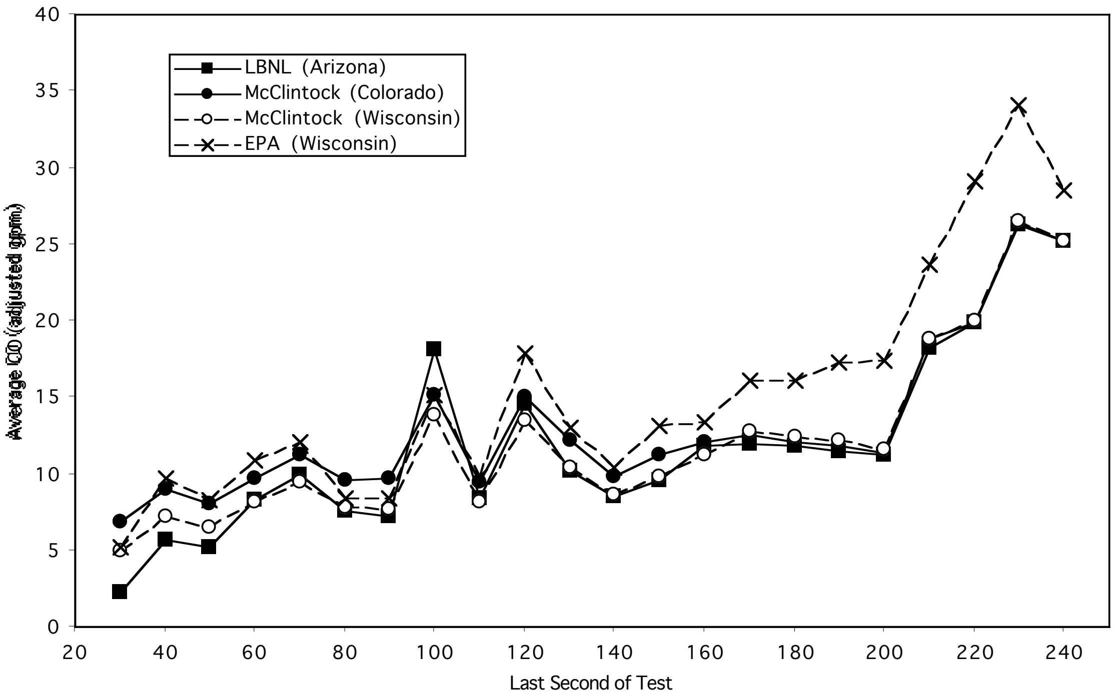
Figures 13 through 15 show average emissions under each prediction method by the last second of the test. Average emissions at each 10-second point only are shown to reduce the complexity of the figure. Each method predicts relatively similar emissions to vehicles that are driven over different portions of the IM240 cycle; the shape of the curves by second of the test are quite similar using each prediction method. Nearly 70% of the cars are passed after only 30 seconds of testing; another 11% are given the full test. The test durations of the remaining 19% of the fleet are fairly evenly distributed over the other 209 seconds of the test.

Note that all but the EPA method converge the further into the test cars are driven; the emissions at second 240 for all but the EPA method are identical. (Since the EPA method should not be applied to cars given the full IM240 test, in the preceding tables and figures the measured values for full IM240s were substituted for the values "predicted" by the EPA method.) The EPA method results in much higher emissions for vehicles driven further into the test than the PM (Wisconsin) method, particularly for CO and NOx. However, as mentioned above, relatively few cars are fast-passed this far into the test.



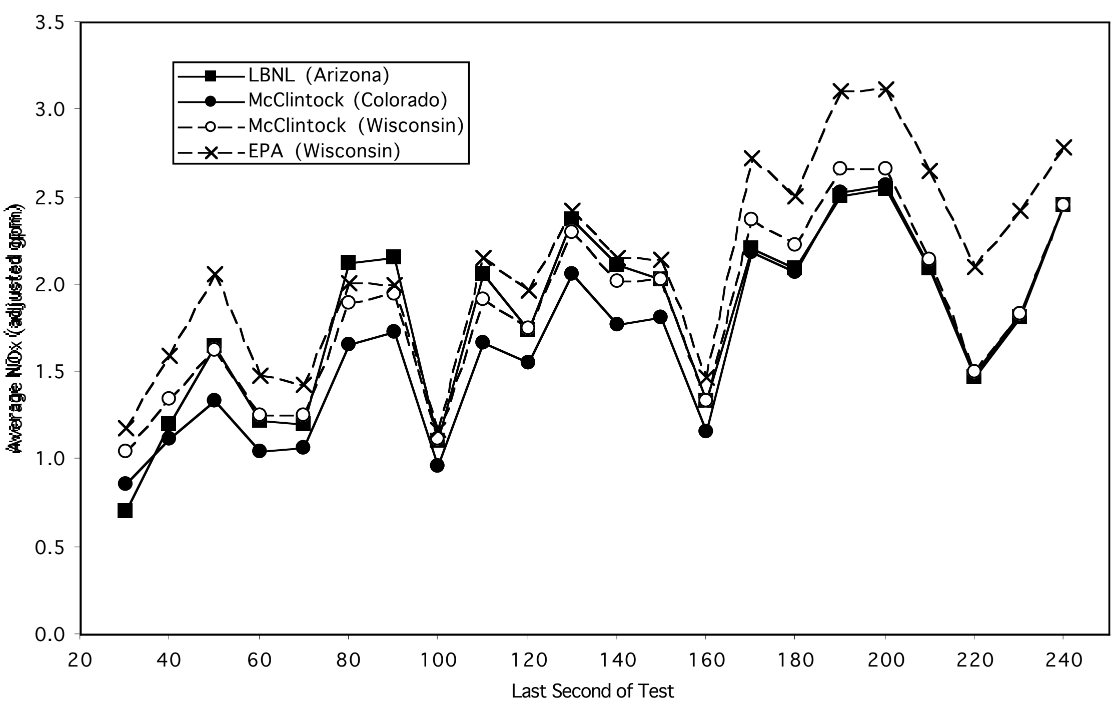
Average CO by Fast Pass Correction Factor and Last Second

MY82-94 Passenger Cars, Wisconsin 1996-97 IM240s



Average NOx by Fast Pass Correction Factor and Last Second

MY82-94 Passenger Cars, Wisconsin 1996-97 IM240s



Evaluation of LBNL Method

Finally, we compare the distribution of emissions from the random sample of vehicles given the full IM240 test in Arizona in 1996, with the adjusted emissions of the vehicles that were not given the full IM240 (i.e. those that were either fast-passed or fast-failed). Figure 16 compares the model year distribution of the cars in each sample, and indicates that the random sample appears to be quite representative of the entire population of vehicles tested under the Arizona I/M program.

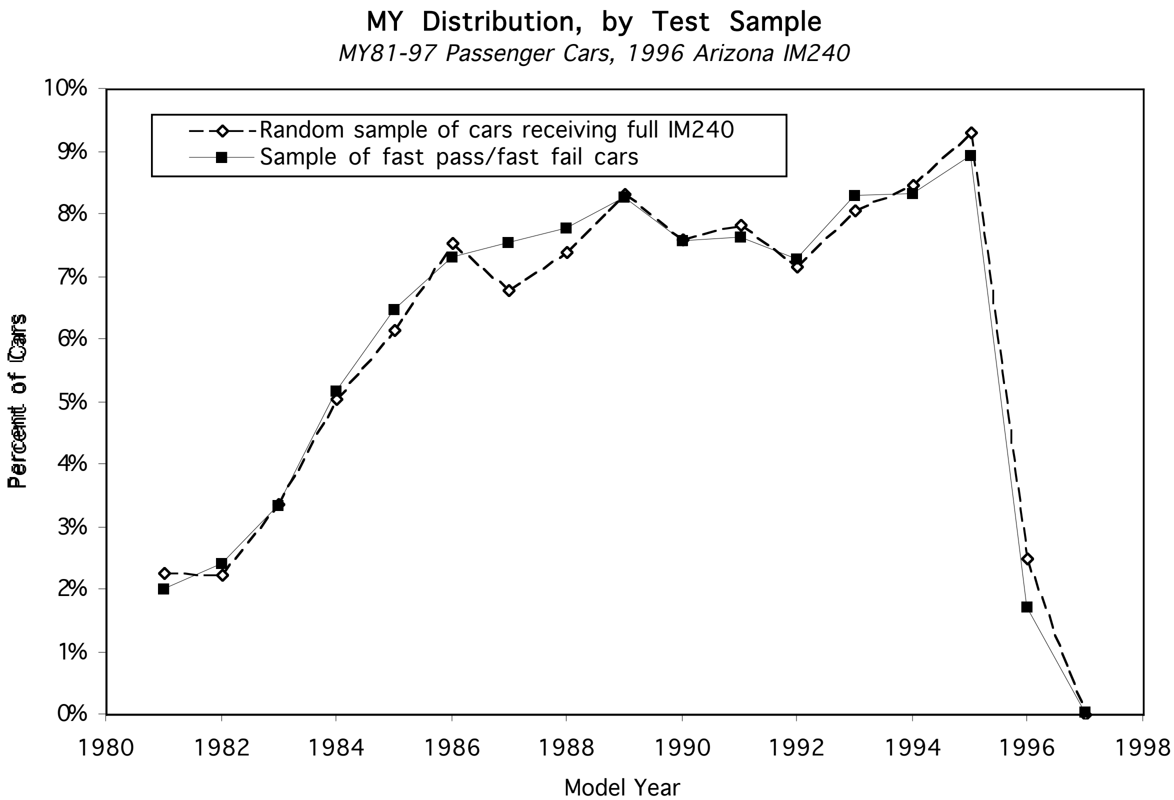


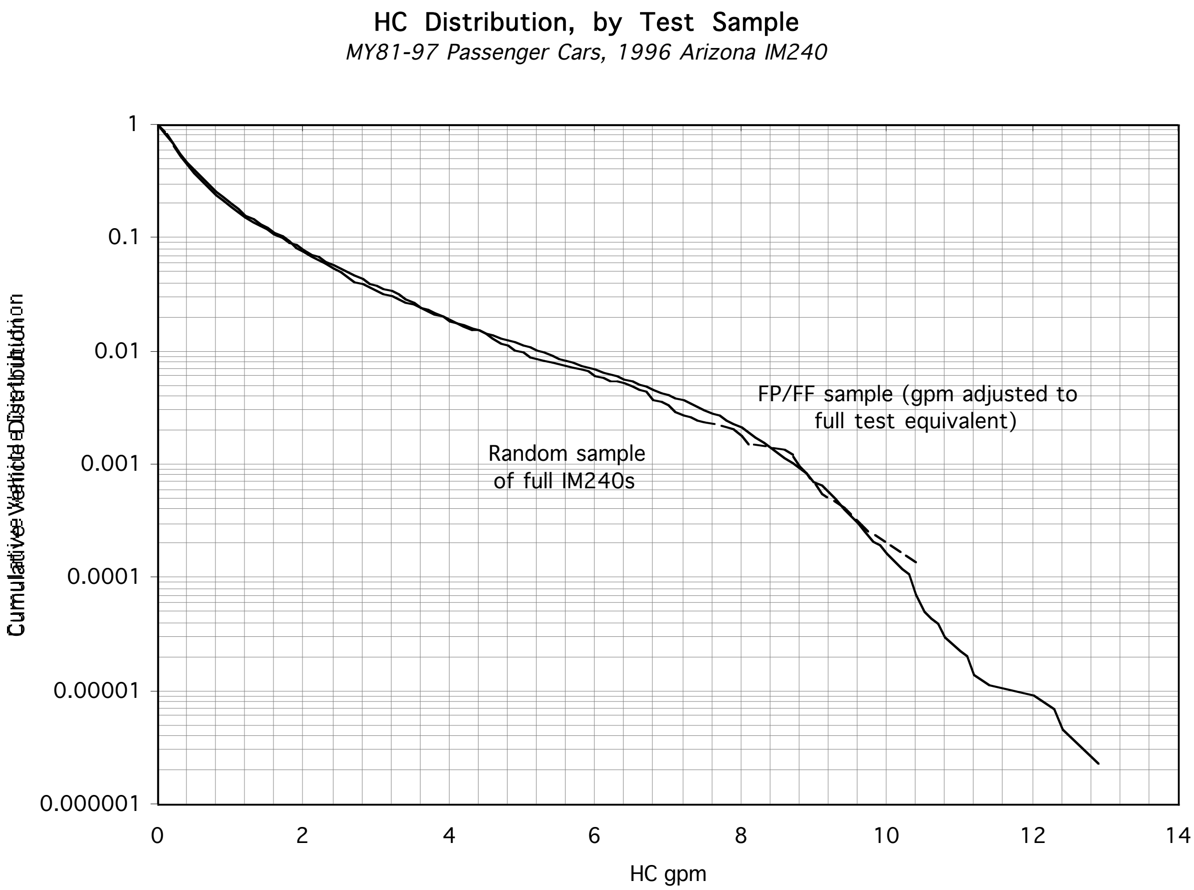
Table 3 compares the measured full test emissions, from the sample of vehicles given the full IM240, with the predicted full test emissions, from the vehicles fast-passing or fast-failing the Arizona IM240. The table indicates that the predicted emissions from the fast-pass/fast-fail vehicles are very similar to those from the random sample of vehicles.

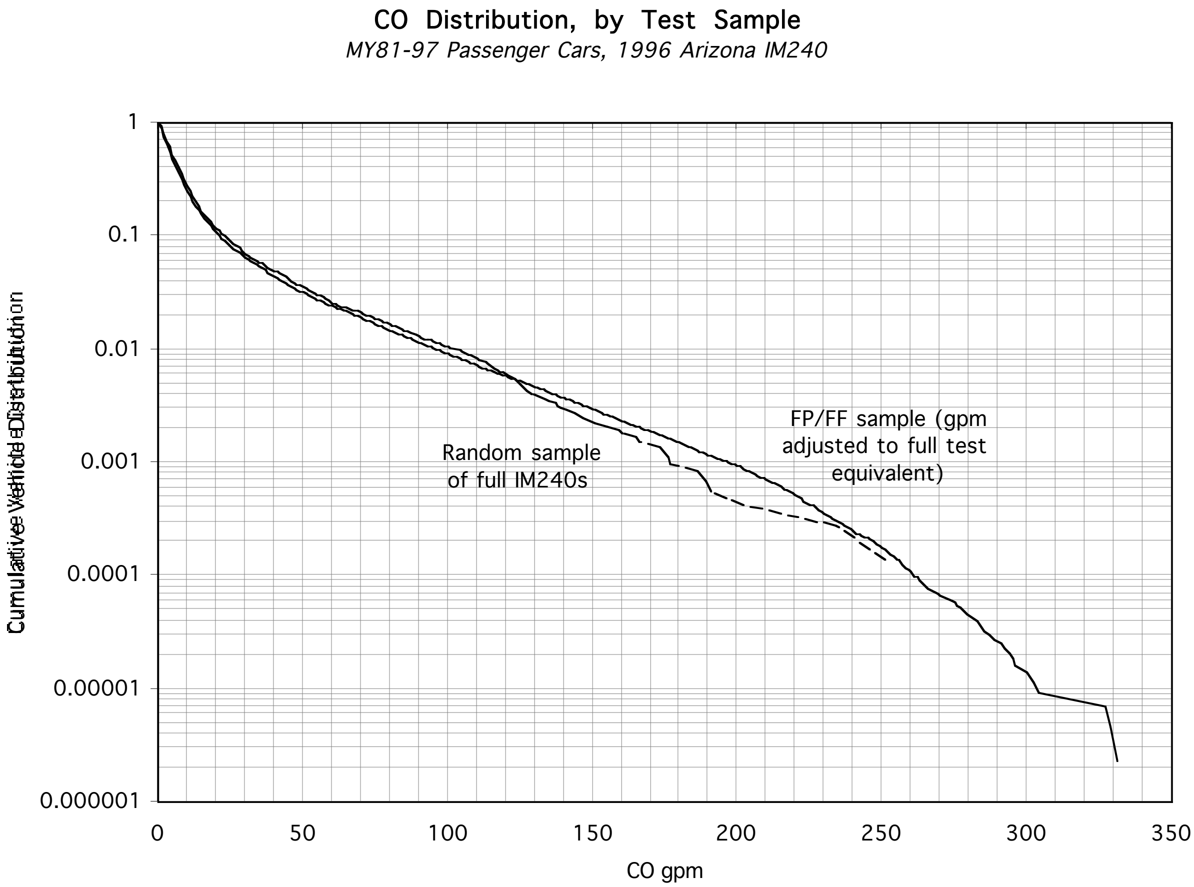
Table 3. Comparison of Measured and Predicted Full Test Emissions, Arizona Random Sample and Fast-Pass/Fast-Fail Tests

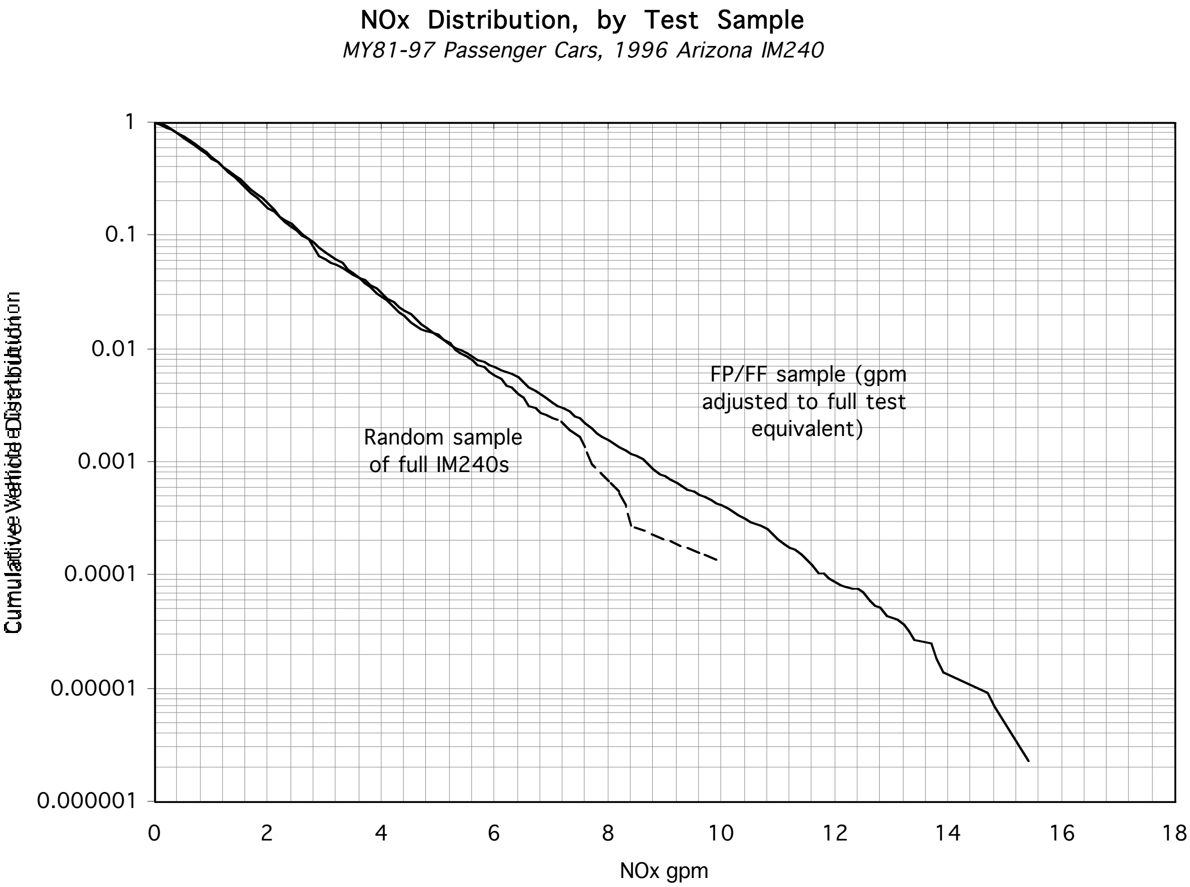
	Average emissions (gpm)			Ratio of FP/FF to full test		
	HC	**CO**	**NOx**	**HC**	**CO**	**NOx**
Random Sample (n=7,209)	0.64	10.3	1.23	1.00	1.00	1.00
Fast Pass/Fast Fail (n=436,160)	0.66	9.5	1.22	1.02	0.93	0.99

Figures 17 through 19 show the emissions distributions of each pollutant by test sample. The figures indicate that the emissions distributions from both the random sample of full tests (dashed line) and the adjusted emissions from the FP/FF tests (solid line) are quite similar.

This similarity contradicts evidence presented earlier that the LBNL method underestimates the emissions of the majority of cars; that is, low emitting cars that pass after only 30 seconds of testing. We have not yet determined possible explanations for this discrepancy.




CO Distribution, by Test Sample
MY81-97 Passenger Cars, 1996 Arizona IM240
Cumulative Vehicle Distribution
1
0.1
0.01
0.001
0.0001
0.00001
0.000001
0
50
100
150
200
250
300
350
CO gpm
Random sample
of full IM240s
FP/FF sample (gpm
adjusted to full test
equivalent)



References

Ando, Amy W., Winston Harrington, and Virginia McConnell. 1998. *Estimating Full IM240 Emissions from Partial Test Results: Evidence from Arizona*. Resources for the Future Discussion Paper 98-24. March.

Enns, Phil, Ed Glover, Penny Carey, Michael Sklar. 1999. *Analysis of Emissions Deterioration Using Ohio and Wisconsin IM240 Data*. Draft report M6.EXH.002. March draft.

McClintock, Peter. 1998. *Further Investigations into Projecting Results for Vehicles that "Fast-Pass" the IM240 Inspection.* Prepared for Colorado Department of Public Health and Environment. January 10 draft.

Wenzel, Tom. 1997. "Analysis of Emissions Deterioration of In-Use Vehicles, Using Arizona IM240 Data." Presented at the SAE Government/Industry Meeting, Washington DC, May 5-7.

APPENDIX I

Seasonal Trends in Vehicle Emissions

Tom Wenzel, Lawrence Berkeley National Laboratory
October 28, 1999

Vehicle emissions as measured by several state I/M programs vary by season. Figure 1 shows the daily average CO of initial IM240 tests of Arizona passenger cars over a three year period (filled circles, left scale). (Emissions of cars that are fast-passed or fast-failed are adjusted to their full IM240 equivalents.) The trend in the maximum daily temperature is also shown (gray lines, right scale). The solid vertical lines denote the calendar years, whereas the dashed vertical lines denote the changes in fuel composition. CO, and HC (Figure 2), are higher in Phoenix in the warmer summer months; on the other hand, NOx shows the opposite seasonal trend, and is higher in winter months (Figure 3). Colorado IM240 data show similar seasonal patterns (Figures 4 through 6).

It is unclear whether the seasonal variation is due to a combination of ambient temperature and changes in fuel composition, or to inadequate conditioning of vehicles prior to testing. Average emissions of MY90 and newer passenger cars that pass their initial I/M test, and therefore would be less likely to be effected by inadequate preconditioning, exhibit the same, albeit muted, seasonal trends in emissions. The seasonal variation in Arizona remote sensing (Figure 7) and loaded idle (Figure 8) CO data appears to mirror that of the Arizona IM240 emissions, suggesting that vehicle conditioning is not the cause of the variation. (The loaded idle data for MY81 and newer passenger cars are taken from the Basic I/M program in Pima County.) However, the seasonal variation in CO (Figure 9) and HC (Figure 10) in the Wisconsin IM240 program and the variation in CO in the Minnesota idle program (Figure 11) are in the opposite direction: CO and HC are higher in winter months. (The extremely high CO values in Minnesota in Figure 11 are likely due to the small number of vehicles tested in these months.) The seasonal NOx trend in Wisconsin, Figure 12, follows that of Arizona and Colorado. A possible cause of the different in the Wisconsin trend from the other states is the use of year-round RFG in the Wisconsin area; however, RFG was introduced in Arizona in late 1997, with no apparent effect on the seasonal variation in emissions. More analysis is needed to better understand these seasonal trends, and why they differ by area.

Figure 1. Daily Average CO, Arizona IM240

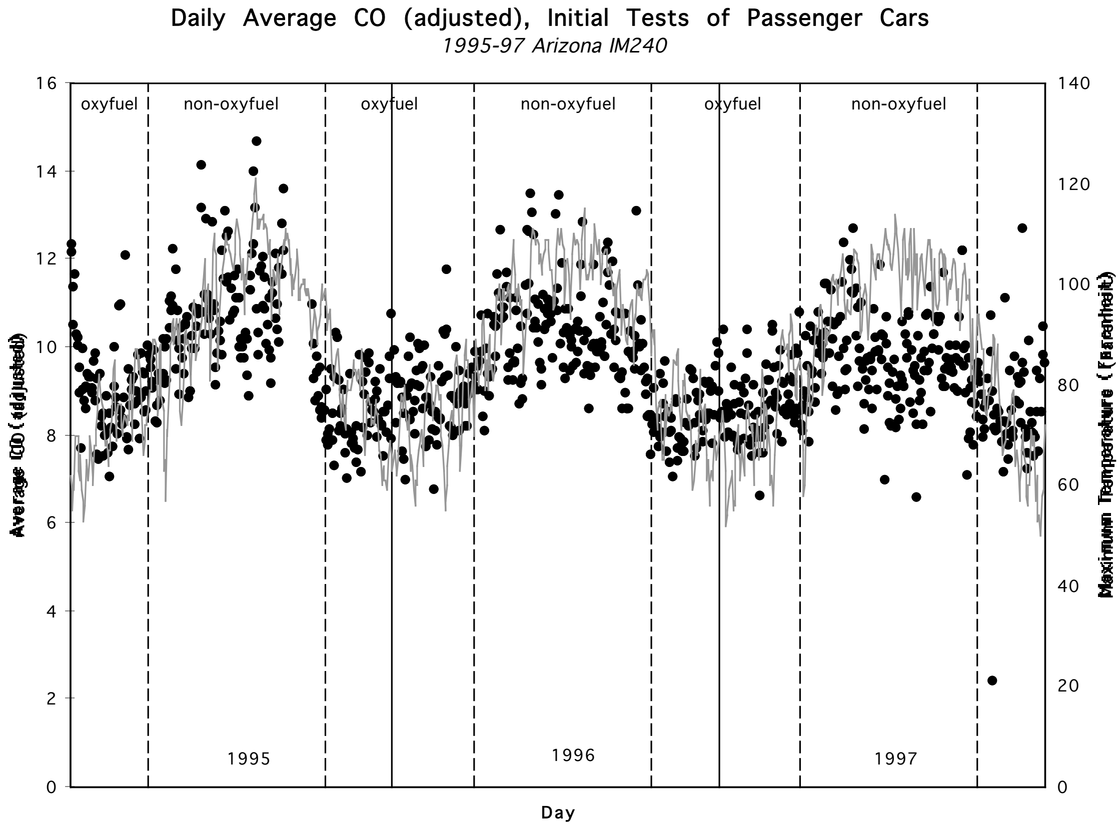


Figure 2. Daily Average HC, Arizona IM240

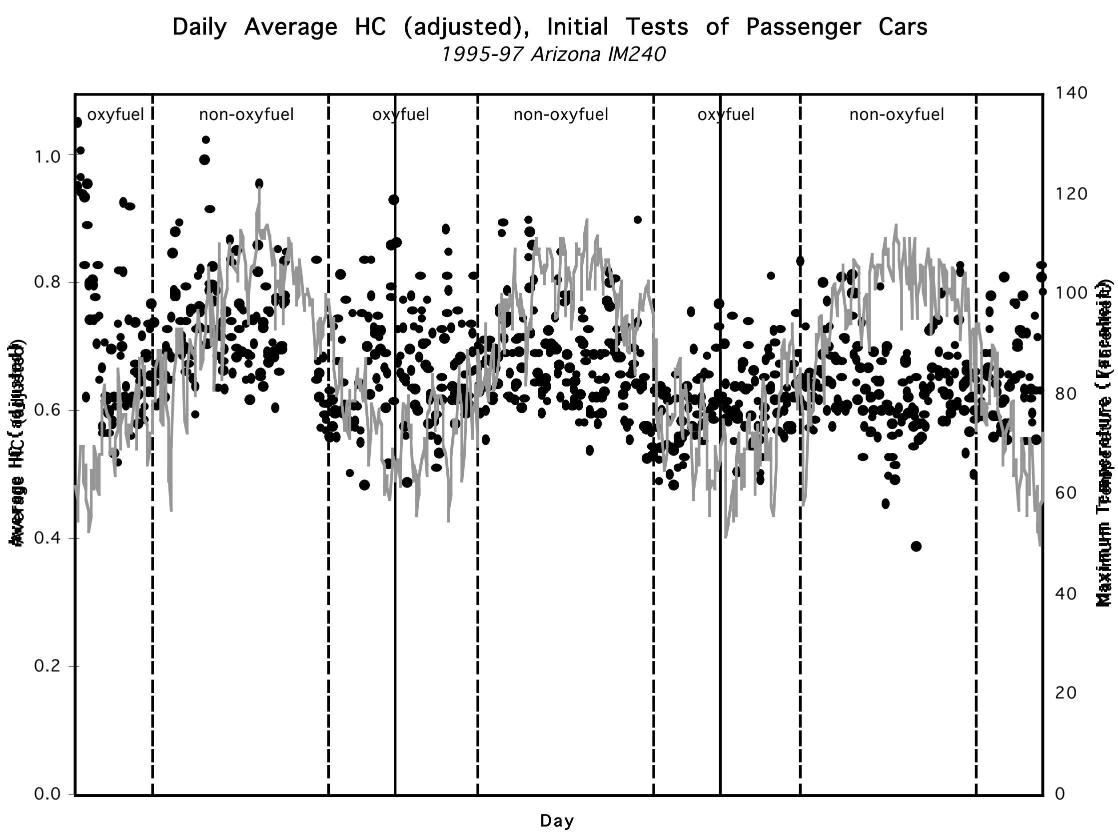


Figure 3. Daily Average NOx, Arizona IM240

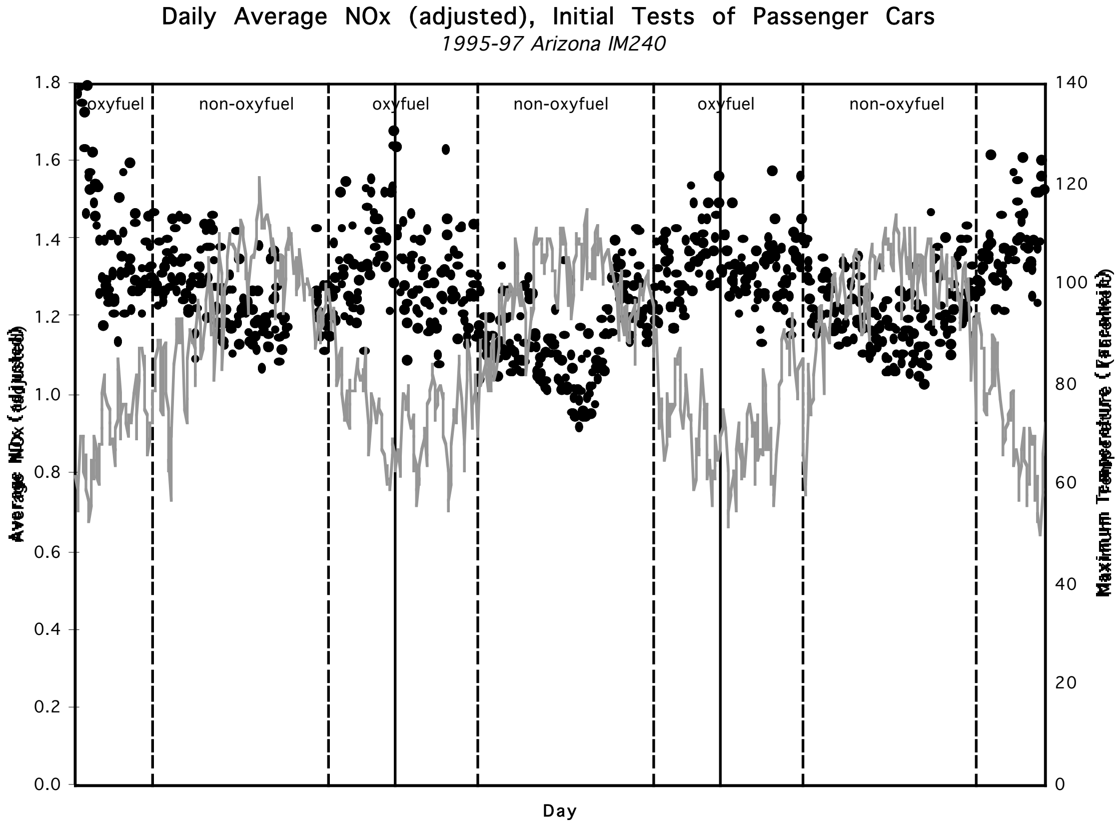


Figure 4. Daily Average CO, Colorado IM240

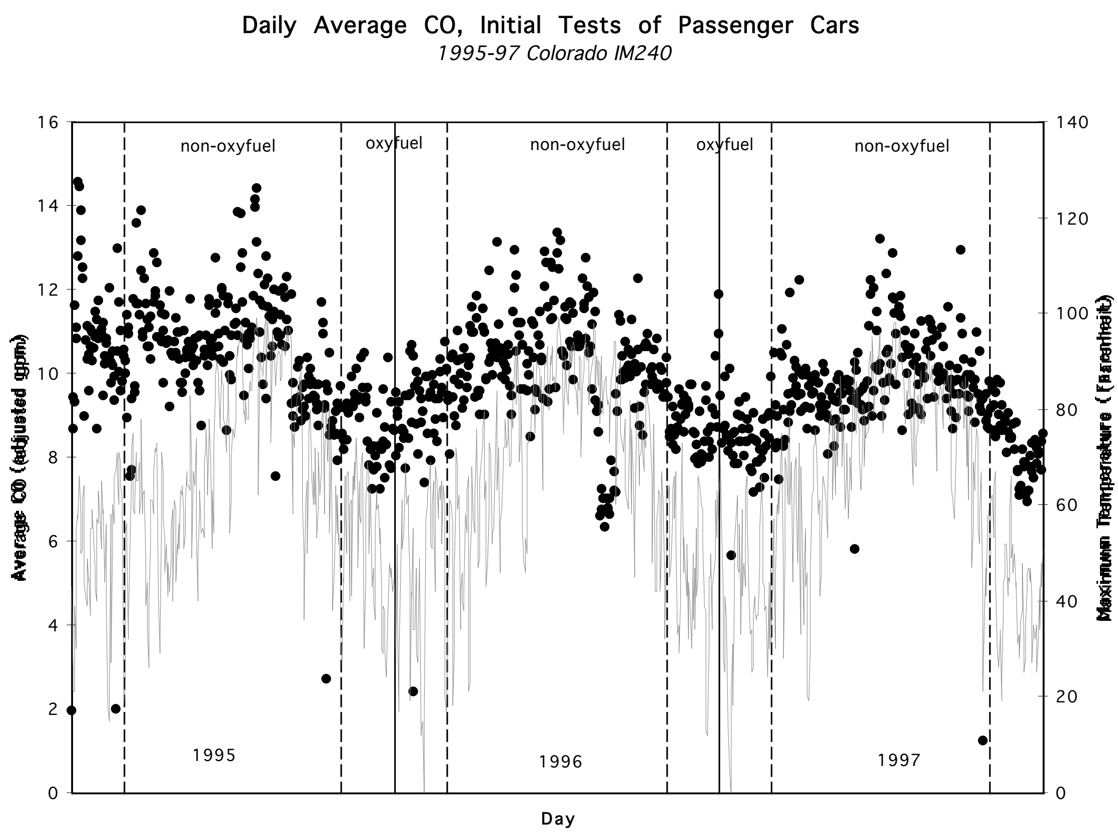


Figure 5. Daily Average HC, Colorado IM240

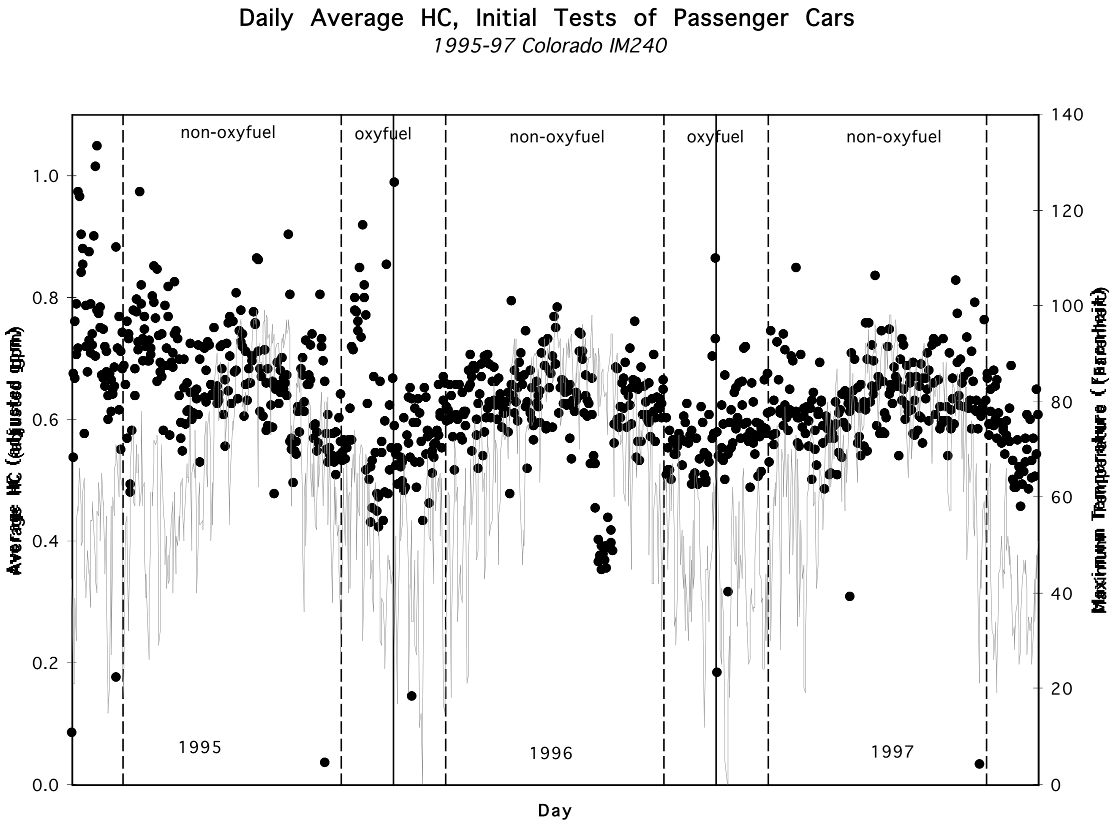


Figure 6. Daily Average NOx, Colorado IM240

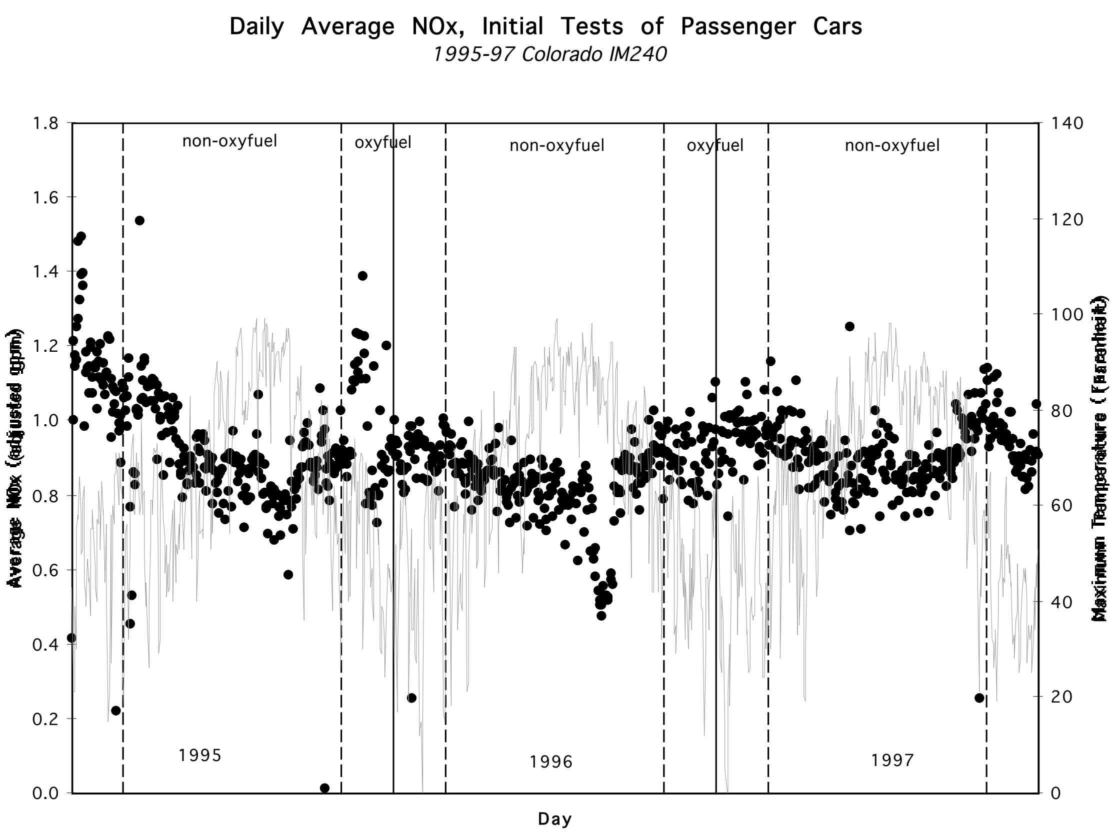


Figure 7. Daily Average CO, Arizona Remote Sensing

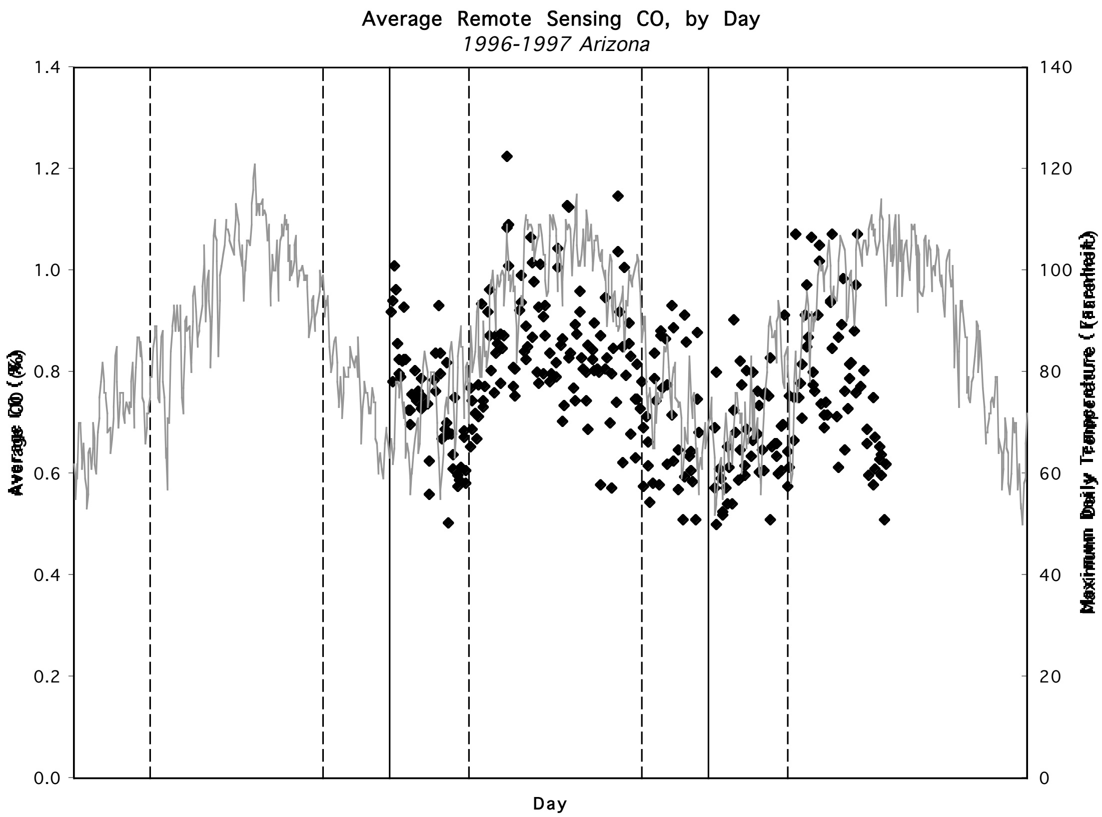


Figure 8. Daily Average CO, Arizona Loaded Idle (Pima County)

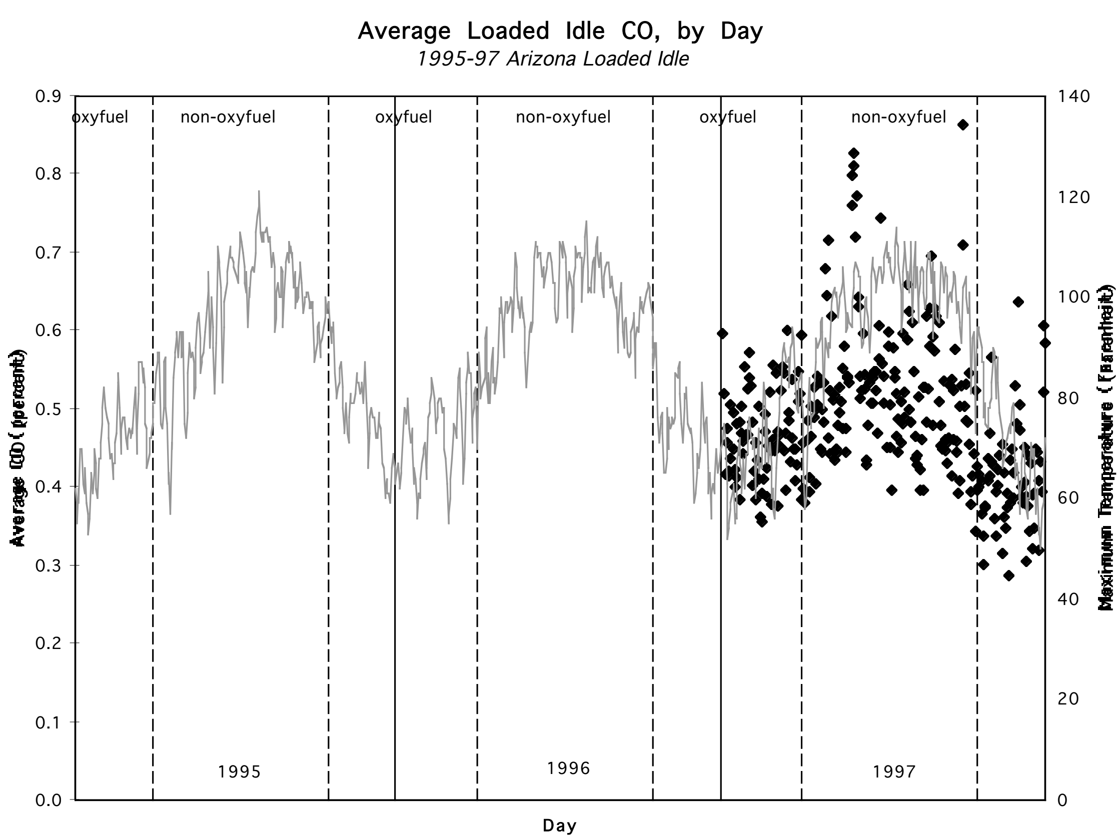


Figure 9. Daily Average CO, Wisconsin IM240

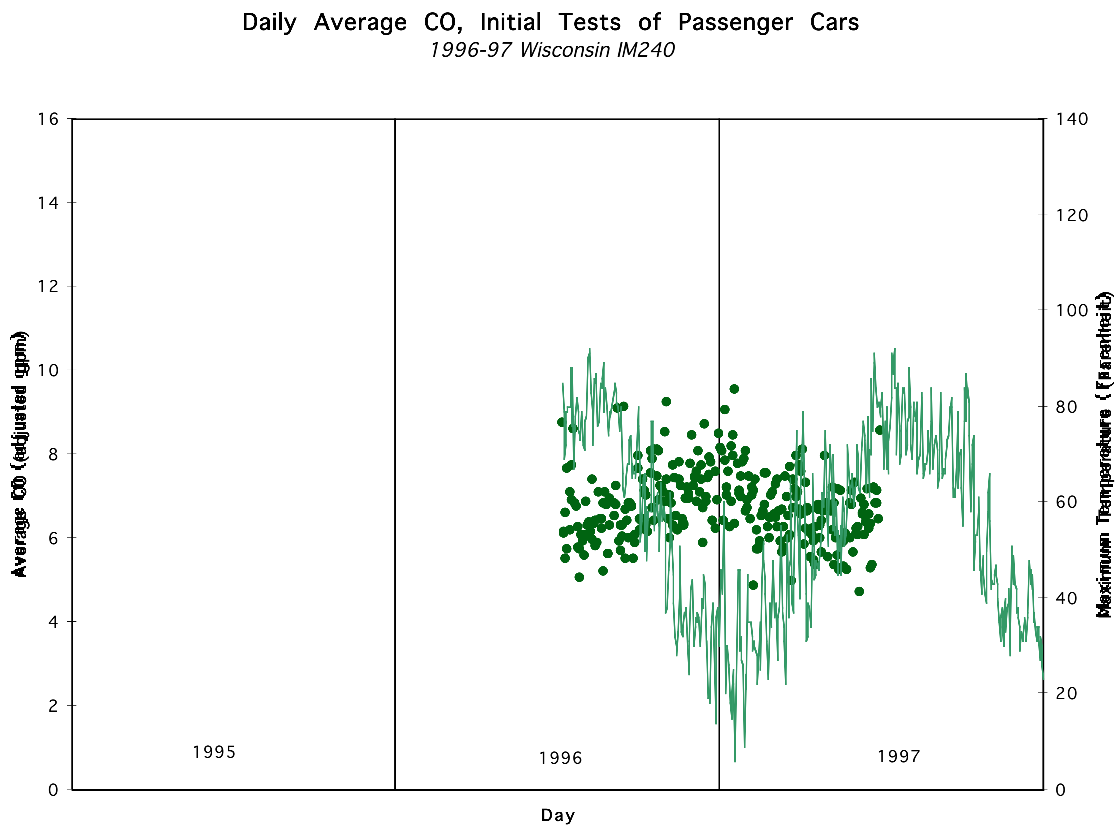


Figure 10. Daily Average HC, Wisconsin IM240

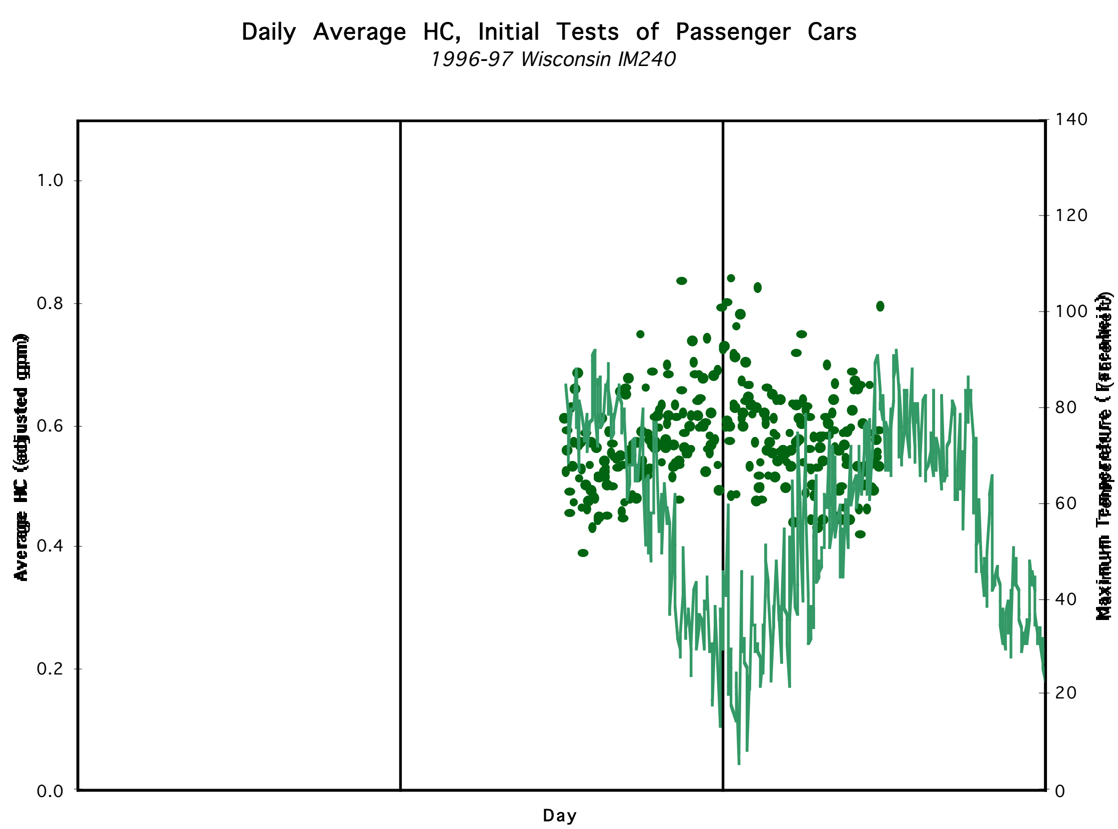


Figure 11. Daily Average CO, Minnesota Idle

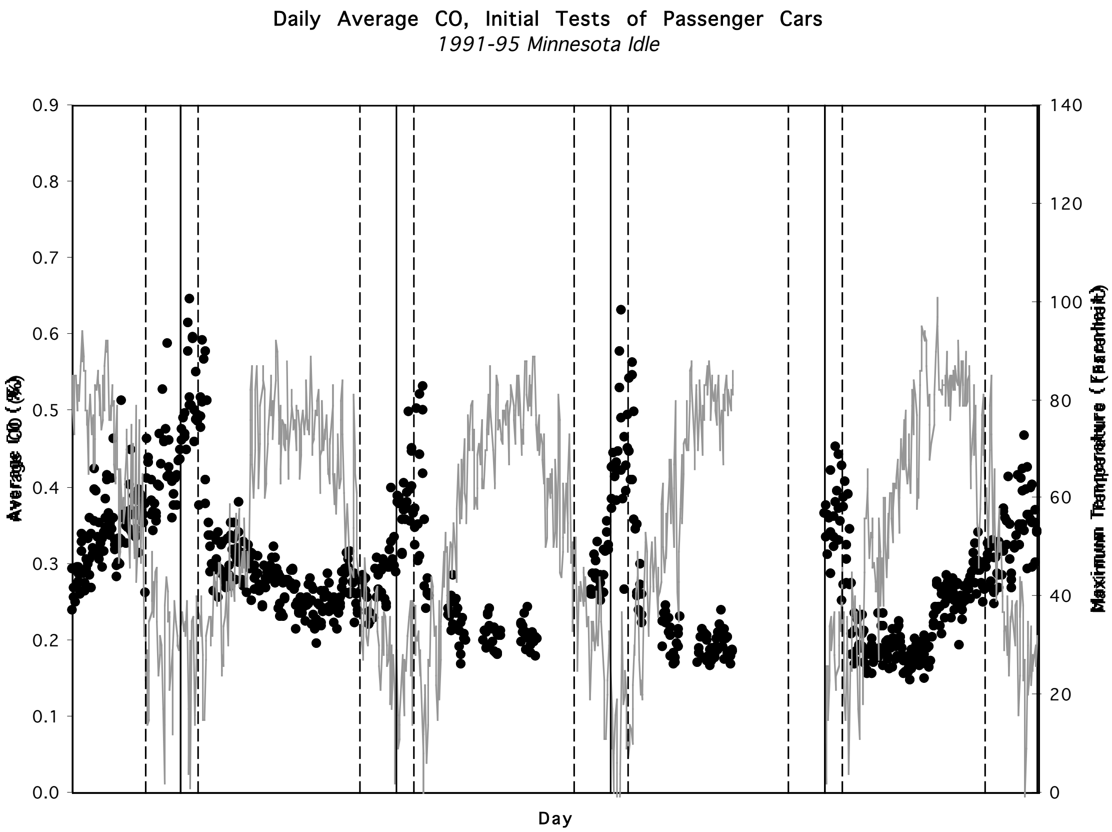
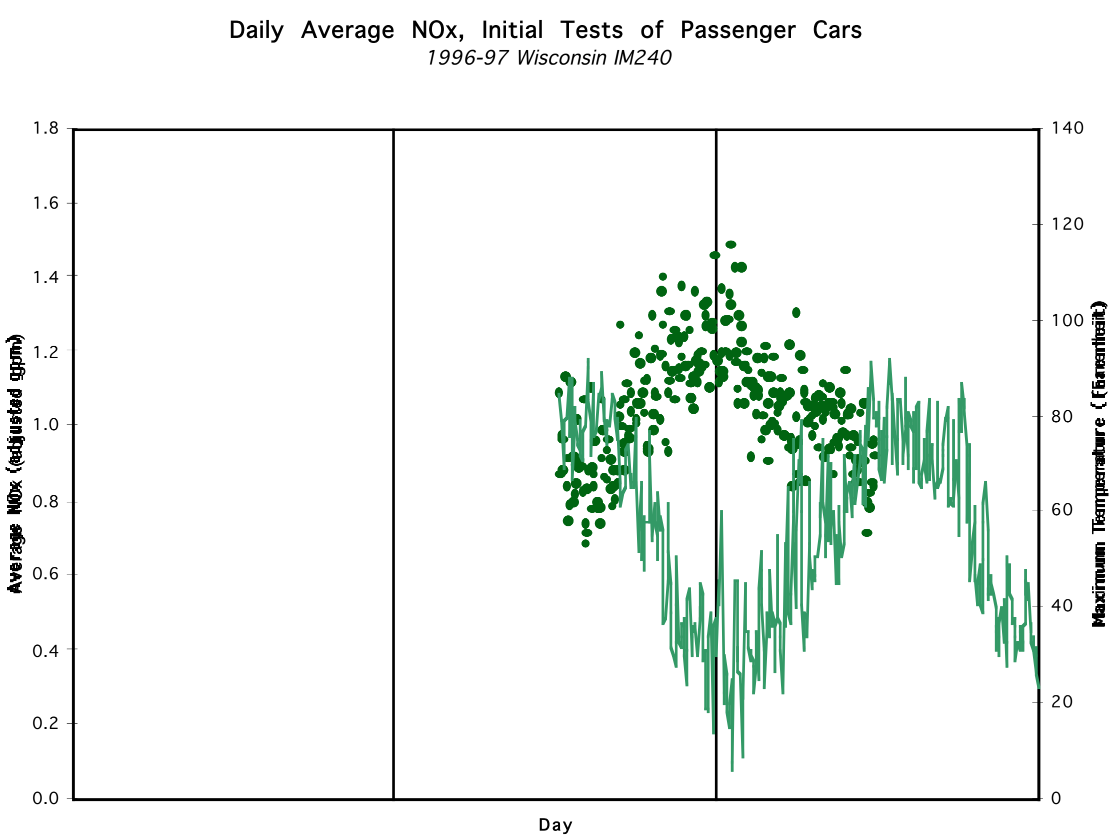


Figure 12. Daily Average NOx, Wisconsin IM240



APPENDIX J

Tracking Vehicles over Time in the Phoenix I/M Program

Tom Wenzel, Lawrence Berkeley National Laboratory

July 27, 1999

14% of the cars tested in 1995 fail the I/M test. Of the cars that initially fail, 41% do not pass a subsequent I/M test by the end of March 1996. (Extending the time in which a vehicle can get a passing test through March 1996 results in 1400 vehicles switching from the No Final Pass category to the Final Pass category, and reduces the No Final Pass rate from 44% of all Initial Fails to 41%.) Approximately 4% of all Initial Fail vehicles are waivered, so about 37% of the Initial Fails are not accounted for. (This waiver rate is from AZ DEQ, apparently based on analysis of the 2% random sample of vehicles.)

41% of the cars tested in 1995 are not tested in 1997. This ranges from 38% of the cars initially passed in 1995, to 72% of the cars not passing by the end of March 1996. (That is, 62% of the cars that initially passed in 1995 were tested in 1997, while 28% of the cars that had no passing test were tested again in 1997). The fraction of cars tested in 1995 that were not tested in 1997 ranges from 61% of the MY81 cars to 35% of the MY94 cars. So the fleet of vehicles not tested in 1997 is older, and has higher average emissions by MY, than the fleet tested in both years.

Why are so many vehicles not tested two years later? Vehicles do not change their I/M test cycle if they are resold, or their registration lapses; the only way a vehicle's test cycle can change is if it is re-registered out of the state and then applies for registration in Arizona, an unlikely occurrence. Apparently Gordon-Darby tracked vehicles that were initially tested in 1996, and found that about 35% did not report for testing in 1998. They attribute this attrition to vehicles relocating out of the I/M area, rather than them avoiding the I/M program. Analysis of remote sensing data, however, indicates that many of these vehicles are still being driven in the Phoenix area. 12% of the fleet tested in both 1995 and 1997 were measured by remote sensors at least 2 years after their initial 1995 I/M test. 5% of the fleet tested in 1995 only were similarly measured by remote sensors. The ratio of the two percentages suggests that as much as 40% of the cars not reporting for testing in 1997 were still being driven in the Phoenix area in 1997.

This technique can also be used to determine what fraction of the vehicles never passing in 1995 are still being driven in the Phoenix area. Remember that 28% of the cars that had no passing test in 1995 were tested again in 1997. Of these cars, 8% were seen by remote sensors at least 2 years after their initial 1995 I/M test. In contrast, 2% of the cars that never passed 1995 testing, and did not report for testing in 1997, were seen by remote sensors at least 2 years after their initial 1995 I/M test. Again, the ratio of the two percentages suggests that 25% of the cars that never passed 1995 testing, and did not report for 1997 testing, were being driven in the Phoenix area in 1997. (This fraction appears to be higher, 40%, for LDTs.)

Of the 5,347 cars that were never passed in 1995 but returned for testing in 1997 (28% of the initial fails in the fleet tested in 1995 and 1997), 65% failed their initial test in 1997. In the Phoenix program, vehicles that do not receive a passing result within 5 months of initial testing have their next test coded as an initial test. Of the 5,347 cars that did not pass through March of 1996, 656 (12%) had second initial tests in 1996; of these, 351 eventually passed in 1996 (54% of those with second initial tests in 1996, 6% of all that did not pass through March of 1996).

Of the cars that initially failed for any pollutant in 1995, one-half failed initial testing in 1997. The percent of repeat failures is much higher for older cars; the percent ranges from 54% for MY81 cars to under 5% for MY94 cars. One-half of the cars that failed initial testing in both 1995 and 1997 failed for the <u>same combination of pollutants</u> in each year. The percent of same type of failure is higher for newer cars; the percent ranges from 44% for MY81 cars to 62% for MY93 cars.

Of the cars tested in both 1995 and 1997, emissions <u>increases</u> between the two test years (due to emissions deterioration in properly functioning vehicles and emissions control malfunction in a relatively small number of vehicles) are greater (26% for HC, 39% for CO, 17% for NOx) than the emissions <u>reductions</u> between initial and final tests in 1995 (12% for HC, 15% for CO, 8% for NOx).

50% of the cars tested in 1997 were not tested in 1995 (223,000). These cars include 52,000 MY94 and older out of state cars newly registered in Arizona (23%), 40,000 MY95 cars exempted from testing in 1995 (18%), and 17,000 MY96 and newer cars voluntarily tested (8%). The remaining 114,000 cars (51%) are voluntary tests, or second initial tests of vehicles that never passed initial testing in 1996, or high emitters flagged by RSD for unscheduled I/M testing. The 1997 I/M results of this "migrating in" fleet is similar to the I/M results of the entire 1995 fleet, including the cars that are exported out of the area after 1995. That is, 15% fail initial testing, and 42% of initial failures do not receive a passing I/M test by the end of 1997. Of the cars that are tested in both 1995 and 1997, only 10% fail initial I/M testing, and 29% of these never receive a passing I/M test, even though this fleet is substantially older than the fleet first tested in 1997. The 1995 "migrating out" fleet has higher average emissions by MY and I/M result than the fleet tested in 1995 and 1997. The average emissions by MY and I/M result of the 1997 "migrating in" fleet is almost identical to those of the 1995 "migrating out" fleet.

Table 1 shows average emissions of the <u>passenger car</u> fleet tested in both 1995 and 1997, by initial and final I/M test in each year. The values are <u>not</u> weighted by annual VMT. Table 2 shows the percent change in emissions in each time period. The first row shows the initial reduction in emissions due to the 1995 I/M cycle. The second row shows the increase in emissions between the final I/M test in 1995 and the initial test in 1997, on the same vehicles. The increase is made up of three factors: insufficient repair of vehicles that failed in 1995; emissions malfunctions of vehicles that passed in 1995; and emissions deterioration due to two years of vehicle aging. The two-year increase in emissions between I/M cycles is the same or greater than the initial reduction due to the program. The third row shows the initial emission reduction from the I/M program in 1997. The last row shows the cumulative effect of two I/M cycles, by comparing the initial emissions in 1995 with the final emissions in 1997. For the car fleet that is tested in both 1995 and 1997, the effectiveness over two cycles of the I/M program is only 6% for HC, 3% for CO, and 1% for NOx. The reductions for the LDT fleet are 7% for HC, 2% for CO, and 0.1% for NOx.

Table 1. Unweighted Fleet Emissions (grams per mile), by I/M Test and Year, All Cars Tested in Both 1995 and 1997

	HC (gpm)	CO (gpm)	NOx (gpm)
1995 initial I/M test	0.57	7.7	1.25
1995 final I/M test	0.50	6.5	1.15
1997 initial I/M test	0.63	9.1	1.34
1997 final I/M test	0.54	7.5	1.24

Table 2. Percent Change in Unweighted Fleet Emissions, All Cars Tested in Both 1995 and 1997

	HC	CO	NOx
Effect of 1995 I/M program (1995 final divided by 1995 initial)	-12%	-15%	-8%
Effect of 2 years of deterioration (1997 initial divided by 1995 final)	26%	39%	17%
Effect of 1997 I/M program (1997 final divided by 1997 initial)	-15%	-18%	-8%
Cumulative effect of two I/M cycles (1997 final divided by 1995 initial)	-6%	-3%	-1%

The percentage changes are slightly different when vehicle emissions, expressed as tons per day, are weighted by annual vehicle miles traveled (using MOBILE6 annual VMT by vehicle type and model year), as indicated in Tables 3 and 4.

Table 3. Fleet Emissions Weighted by Annual VMT (tons per day), by I/M Test and Year, All Cars Tested in Both 1995 and 1997

	HC (tpd)	CO (tpd)	NOx (tpd)
Effect of 1995 I/M program (1995 final divided by 1995 initial)	4.11	55.40	9.29
Effect of 2 years of deterioration (1997 initial divided by 1995 final)	3.66	47.99	8.23
Effect of 1997 I/M program (1997 final divided by 1997 iniital)	4.50	65.17	10.09
Cumulative effect of two I/M cycles (1997 final divided by 1995 initial)	3.88	54.37	9.06

Table 4. Percent Change in Weighted Fleet Emissions, All Cars Tested in Both 1995 and 1997

	HC	CO	NOx
Effect of 1995 I/M program (1995 final divided by 1995 initial)	-11%	-13%	-11%
Effect of 2 years of deterioration (1997 initial divided by 1995 final)	23%	36%	23%
Effect of 1997 I/M program (1997 final divided by 1997 iniital)	-14%	-17%	-10%
Cumulative effect of two I/M cycles (1997 final divided by 1995 initial)	-6%	-2%	-2%

Figure 1 shows the weighted data in Tables 3 and 4. The figure indicates that the initial reduction from the first cycle of the Enhanced program is equivalent to the reduction from future cycles. (The Phoenix area had a Basic program in place prior to implementation of the Enhanced program, which could have muted the first year effect of the Enhanced program. On the other hand, subsequent cycles of the Enhanced program achieve roughly the same emissions reduction as the first cycle.) In addition, the effect of two years of emissions deterioration outweigh the effect of the I/M program; the after program emissions in 1997 are substantially higher than the after program emissions in 1995.

These curves present an optimistic estimate of the effect of the I/M program. This is because the analysis is limited to vehicles reporting for testing in both 1995 and 1997. About 40% of the cars reporting in 1995 did not report for testing in 1997; a fraction of these cars are surely still being driven in the I/M area (we estimate 40% based on remote sensing data, discussed above). The curves are based on 28% of the no final pass vehicles in 1995 returning for testing in 1997 (as observed in the I/M data, discussed above); this percentage is comparable to the 25% estimate of No Final Pass vehicles still being driven in the I/M area, derived from RSD data.

Figures 2 and 3 decompose Figure 1 into subfleets, based on each vehicle's 1995 I/M result. Figure 2 shows that the Initial Fail/Final Pass vehicle emissions are reduced dramatically, primarily due to repairs or adjustments. But from 1995 to 1997 their emissions nearly double, due to ineffective repair or adjustments that result in a passing test but no real emissions reduction. About half of the initial reduction, as measured by comparing first and last 1995 I/M tests, is lost. Emissions from the No Final Pass and No Second Test cars <u>decrease</u> between 1995 and 1997. A possible cause is that repairs or adjustments were made to these vehicles after their last 1995 I/M test that reduced their emissions. Another possibility is the emissions variability of high emitting vehicles (possibly due to intermittent malfunction of emissions controls); if tested again, many of these high emitters would exhibit lower emissions. As noted above, 65% of the cars that did not pass 1995 I/M testing failed their initial 1997 I/M test. Figure 3 shows the emissions of cars that initially passed in 1995.

Figure 1. Fleet Emissions over Two I/M Cycles
Passenger Cars tested in both 1995 and 1997, Arizona IM240

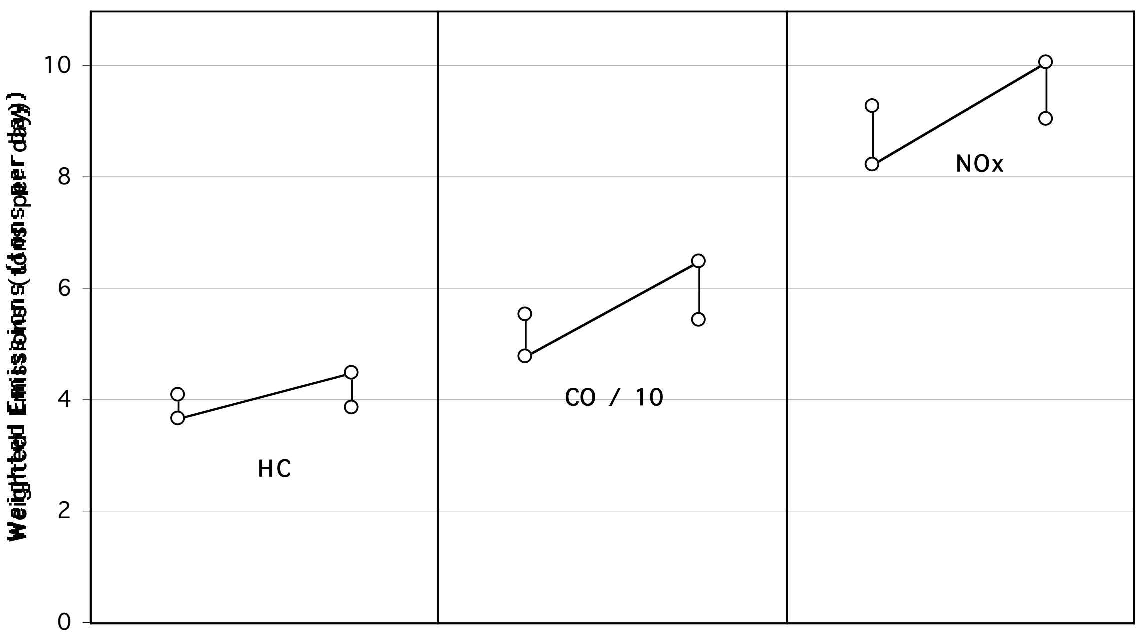


Figure 2. Fleet Emissions over Two I/M Cycles
Passenger Cars tested in both 1995 and 1997, Arizona IM240

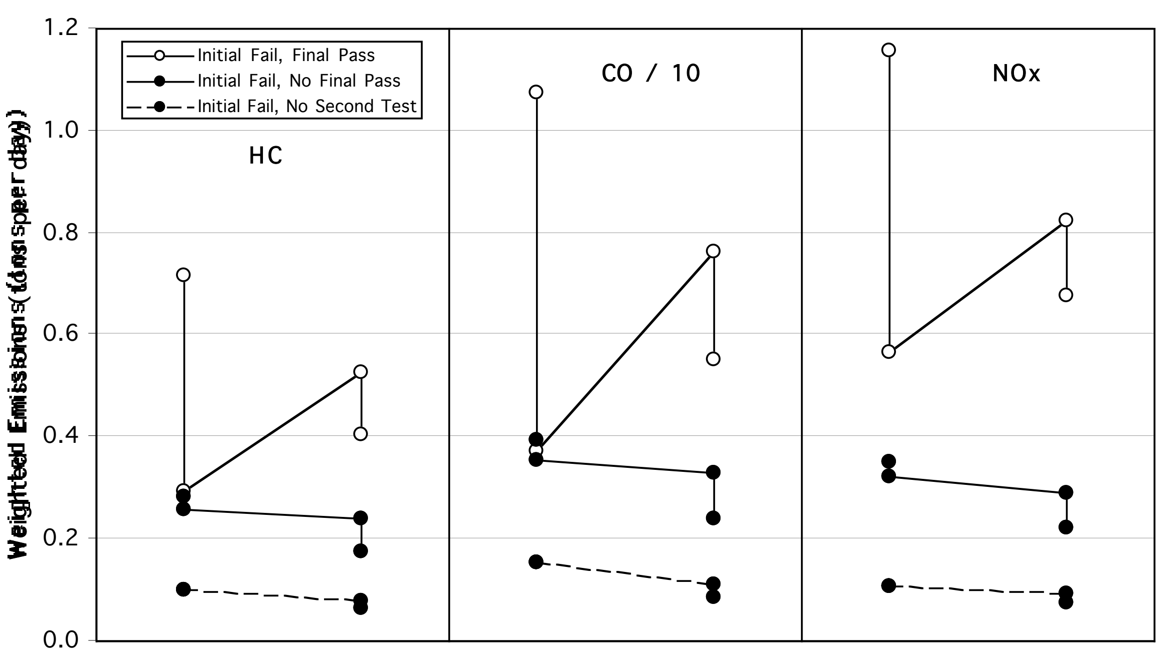
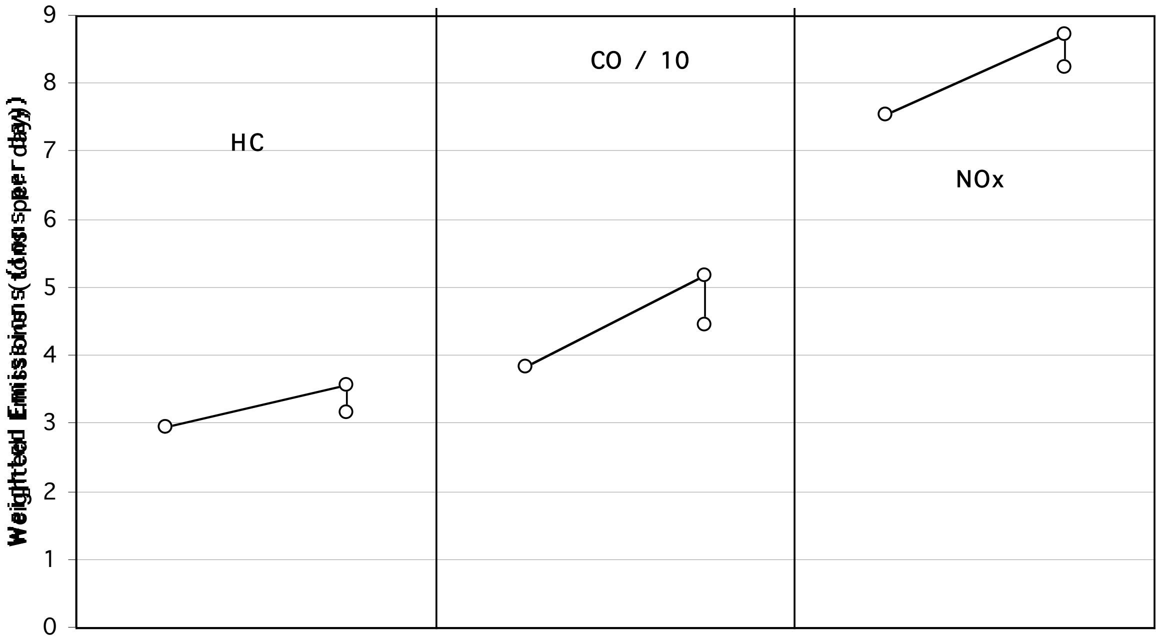


Figure 3. Fleet Emissions over Two I/M Cycles

Passenger Cars tested in both 1995 and 1997, Arizona IM240



APPENDIX K

APPENDIX L

APPENDIX M

APPENDIX N

Comments on EPA's Draft Report *MOBILE6 Inspection/Maintenance Benefits Methodology for 1981 through 1993 Model Year Light Vehicles, M6.IM.001*

Submitted by Tom Wenzel, Lawrence Berkeley National Laboratory
July 27, 1999

1) EPA, page 4: *12. High Emitter Waiver Rate - Selected to be 15 percent of failures, and loosely based on analysis of Arizona and Ohio I/M vehicles.*

Comment: This rate is too high; Arizona and Wisconsin report waiver rates of 4 percent of all failures.

2) EPA, pages 10 and 11: *Line D represents the average emissions of the portion of high emitting vehicles that are identified and repaired because of the I/M process. This line is calculated as a function of vehicle age, and is a percentage (e.g., 150%) of Line B. The portion of the fleet which is identified by I/M will be repaired to a lower level on average. However, this level is not as low on average as the average of the normal vehicles. The justification for this assumption was an analysis of Arizona IM240 before and after repair data collected during 1995 and 1996.*

and

EPA, page 20: *The normal emitter emission level is used as the final after repair emission level if it is larger than the calculated after repair emission.*

Comment: LBNL analysis of Arizona IM240 data finds the same result, that average emissions of seemingly repaired vehicles are not brought down to the level of vehicles that pass their initial test (Wenzel, 1999a). Figures 1 though 3 show the factors necessary to adjust average emissions of initially passing vehicles to the average emissions of vehicles that failed initial but passed final I/M testing in the Arizona program in 1996-97, compared with the factors EPA is proposing for MOBILE 6. In general the factors are quite similar, although EPA's factors are slightly higher, particularly for HC and CO.

3) EPA, pages 12 and 13: in –use deterioration regression coefficients for normal emitter cars.

Comment: The curves for the different age groups of each fuel technology (FI and Carb) are roughly parallel (see Figures 4 through 9). However, all FI categories have higher slopes than all Carb categories. This means that at high mileages FI vehicles have higher emissions than carbureted vehicles. For instance, 1988-93 TBI and PFI HC emissions exceed 1986-89 Carb HC emissions around 80,000 miles (Figure 4). At 150,000 miles, HC emissions are: 0.26 TBI, 0.23 PFI, and 0.20 Carb. This does not make sense; a 1993 PFI is supposed to be much more durable than a 1986 Carb. If anything, the FI vehicles should have smaller slopes, in line with manufacturer claims (and evidence from IM240 data) that newer TBI and PFI technology vehicles have much less emissions deterioration than older FI and Carb technologies. In addition, for cars the CO PFI curve is much higher than the TBI curve. There is anecdotal evidence from some manufacturers that early PFI technology had problems to be worked out; however, this should not affect all years of PFI. The NOx deterioration rates for carbureted light trucks (Figure 9) have a very high intercept, and a very small slope. These rates suggest that these vehicles experience very little emissions deterioration as they accumulate mileage.

4) EPA, page 14: *The high emitter HC emission level for the 1988-93 MY PFI group is also a special case. For this group it was thought that the average high emitter emission level was too low because it caused the average high emitter level to be lower than the normal emitter level at fairly low mileages. It was increased from 1.10 g/mi HC to 1.74 g/mi HC by adding one very high emitting 1987 model year vehicle to the 1988-93 model year PFI group.*

Comment: The rationale for moving a vehicle into this group does not make sense. Based on the coefficients in Table 1a, a 1993 PFI car with 150,000 miles would have HC emissions of 0.16 gpm, which is much less than the 1.1 gpm mean emissions of high emitter cars.

5) EPA, page 14: *An analysis of the Ohio IM240 data was also done to try and estimate the high emitter levels for running LA4 and start emissions. This was done because of the small numbers of high emitters in the EPA and AAMA FTP (running LA4 and Start) data samples. In this analysis, a large sample of Ohio vehicles were segregated into normal and high emitters, and the average high emitter emission levels were determined and compared with the FTP based estimates. They compared favorably. However, the analysis was plagued with uncertainties such as how to separate the normals from the highs when FTP data are not available, the inability to split PFI from TBI in the Ohio IM240 data, …*

Comment: The VIN does not consistently distinguish between different fuel delivery systems; consequently, the only way to determine the fuel delivery system of a particular vehicle is to visually inspect the vehicle. VIN decoders attempt to identify the type of fuel system for individual vehicle models; however, such decoders are subject to error (for instance, one such decoder identifies late 1980s Hyundai Excels as fuel injected, whereas they used carburetors during those years). In addition, VIN decoders typically do not distinguish among different fuel injection systems. EPA should review certification records, or survey manufacturers, to determine what fuel system was installed in different years of each vehicle model.

6) EPA, page 15: *Because of these problems the Ohio IM240 data were not used to estimate the average high emitter emission levels.*

Comment: There are recognized problems with IM240 data, as EPA notes. However, these data provide emissions information on a huge, relatively unbiased, sample of vehicles. EPA should consider examining multiple years of IM240 data in order to check its conclusions on in-use deterioration based on FTP data. EPA should compare the results from the Ohio IM240 data with those from other state IM240 programs (Arizona, Colorado, and Wisconsin). An analysis of other state IM240 data would include any cumulative effect of basic I/M programs on in-use emissions. However, evidence indicates that basic programs may not have had a dramatic cumulative effect on in-use emissions.

7) EPA, page 19: Table 2c.

Comment: The report should state clearly that the average after repair HC levels in Table 2c exceed the HC cutpoints. Was this really the case? If trained mechanics cannot repair HC emissions to below the IM240 cutpoints, this has important implications for I/M program effectiveness.

8) EPA, pages 22 and 23: Table 2f, Technician Training Emission Effects

Comment: The percent different between the "after repair by Master Tech" and "after repair by Student Tech" should not be considered the effect of technician training on repair effectiveness. Rather, the Master Tech levels should be considered the maximum emissions reduction potential of vehicle repair. EPA should find some other way to simulate the effect of technician training on repair effectiveness.

9) EPA, page 24: *Because no analysis has yet been conducted on data from operating IM240 programs to estimate the after I/M emission level of vehicles which were waived from the requirement to pass the test, an assumed reduction percentage will have to be used, or the individual user will have to provide a value. The default value will be a 20 percent reduction from the high emitter line for all pollutants.*

and

EPA, page 4: *10. Waiver Repair Levels - In MOBILE6, cost waivered I/M failures will get some repair benefit. A value of a 20 percent reduction has been chosen.*

Comment: Gordon-Darby tracks the average emissions levels by model year of vehicles that receive waivers in the Arizona I/M program, from the 2% random sample of vehicles that received a full IM240. These could be compared to the emissions of intial pass vehicles, to determine how much higher emissions from waived vehicles are from emissions of initial pass vehicles. Other IM240 states may have similar data.

10) EPA, page 25: *For the model year groups of 1981-82 and 1983-85 HC and CO emissions, it was found that the base emission factors at higher mileage levels become higher than the average emissions of the high emitters. It occurs because at high mileages the basic emission factors are data extrapolations. However, under the structure of the model, this is not possible, and it implies that the fleet contains more than 100 percent high emitters. To overcome this inconsistency, it was assumed that the average base emission factors could not continue to rise after it reaches the average of the high emitters, and that it would be set to the average of the high emitters. Typically, the cross-over point is between 150,000 and 200,000 miles, and after this point is reached, it is assumed that the percentage of highs in the fleet for this model year group / technology is 100 percent. This flattening of the emission factor line at very high mileages is consistent with some remote sensing studies. A physical explanation would be that*

while some surviving vehicles continue to deteriorate, the worst emitters are progressively scrapped out of the fleet in the high mileage range.

Comment: This assumption essentially assumes that no early-1980s vehicles are normal emitters. In the Arizona program, about half of the 1981-85 cars with odometers between 150,000 and 200,000 miles passed tailpipe testing, using start-up cutpoints. The normal vehicle deterioration rate should be adjusted, perhaps on the basis of analysis of IM240 data, so that the average emissions of normal vehicles at high mileage are less than the average emissions of high emitters.

11) EPA, page 27: *In the MOBILE6 model, the non-compliant vehicles will be represented as a fraction of the identified high emitters that did not pass or receive a cost waiver. A default value of 15 percent will be built into the model for the non-compliance rate…As an approximation, it is assumed that the 15 percent non-compliance rate (from above) includes the effect of high emitters which did not show up for their first test.*

and

EPA, page 4: *11. High Emitter Non-Compliance Rate - Set to a default value of 15 percent. MOBILE6 will offer users the ability to enter alternative values. This is a generous default which is based on extensive analysis of Arizona and Ohio I/M vehicles. The analysis suggested higher rates (> 20 percent). It also includes high emitters which do not show up for the initial I/M test.*

Comment: LBNL's analysis of Arizona IM240 data indicates that 30 to 40% of vehicles that fail their initial I/M test do not eventually receive a passing result (Wenzel, 1999a; Wenzel, 1999b). Even accounting for the 4% of all vehicles that fail initial testing that receive a waiver, the EPA estimate of 15% appears to be low. Analysis of remote sensing data indicate that 25% of these vehicles are still being operated in the I/M area more than 2 years after their last I/M test (Wenzel, 1999b). A significant portion of vehicles may not be reporting even for initial I/M testing. Analysis of two years of IM240 data in Arizona indicates that 40% of the vehicles tested in 1995 did not report for testing in the next I/M cycle (1997). Remote sensing data indicate that about 40% of these vehicles were still being driven in the I/M area at least two years after their 1995 I/M test (Wenzel, 1999b). This suggests that 16% of <u>all vehicles</u> avoided initial I/M testing during the second cycle of the Enhanced Arizona program.

The default noncompliance rate in the MOBILE6 model should be a conservative estimate (that is, a higher rate should be the default rate), in order to encourage states to gather data in support of a more accurate noncompliance rate.

12) EPA, page 31: *Existing evidence suggests that the type of problems which cause I/M failures can re-occur as often in the repaired vehicles as they do in the unrepaired fleet. Thus, it is assumed that the fleet, after repair, will have the same emission deterioration as before repairs.*

Comment: Nearly 40% of the vehicles that failed initial testing, and passed final testing, in 1995 failed initial testing in 1997. This suggests that repaired vehicles have a <u>greater</u> probability of failing subsequent I/M testing than initially passed vehicles. (The repeat failure rate increases by vehicle age, with over 40% of MY81 cars failing their initial 1997 I/M test, and 15% MY94 cars failing their initial 1997 test. Of the repeat failures, about half failed for the same combination of pollutants in each test year.) (Wenzel, 1997c) An analysis of the Colorado IM240 program found a similar high percentage of vehicles repeatedly failing their initial I/M test in subsequent years (ENVIRON, 1998). Similar analyses of multiple test cycles of other state I/M programs can be performed to determine the repeat failure rate in other states.

13) EPA, Figure 2A, Annual I/M Credits Sawtooth.

Comment: The figure suggests that the initial emissions reduction after introduction of an I/M program is dramatically larger than the emissions reductions of subsequent I/M cycles. The figure also suggests that the effect of the I/M program is to offset nearly all of the increase in emissions due to deterioration from increasing mileage. An analysis of the cars that were tested in both 1995 and 1997 of the Arizona IM240 program suggests a different picture of the effect of the I/M program (see Figure 10, below, and attached memo). Here the initial reduction from the first cycle of the Enhanced program is equivalent to the reduction from future cycles. (The Phoenix area had a Basic program in place prior to implementation of the Enhanced program, which could have muted the effect of the Enhanced program. On the other hand, subsequent cycles of the Enhanced program achieve roughly the same emissions reduction as the first cycle.) In addition, the effect of two years of emissions deterioration outweigh the effect of emissions repairs, such that the after repair emissions in 1997 are substantially higher than the after repair emissions in 1995. Figure 2A may be illustrative only; EPA should determine if the results obtained from MOBILE6 look like those in Figure 10, and, if not, make adjustments to the model assumptions. Again, tracking individual vehicles over multiple I/M cycles can be performed for I/M programs in other states.

14) EPA, page 64: *Neither the Idle Test or the 2500RPM/Idle test will produce NOx benefits or NOx "Dis-benefits" for MOBILE6. In comparison, MOBILE5 contained NOx "Dis-benefits" if an Idle or 2500RPM Idle test were performed.*

Comment: The assumption in MOBILE5 that a basic I/M program that tested for HC and CO only would result in an increase in NOx emissions is a good one. Studies of repair effectiveness typically indicate that repairing for HC and CO only can lead to increases in NOx emissions. EPA should describe on what basis they are recommending removal of NOx disbenefits for programs not testing for NOx.

References

ENVIRON, et al. 1998. Performance Audit of the Colorado AIR Program, prepared for the Office of the State Auditor, State of Colorado, March.

Wenzel, Tom. 1999a. "Evaluation of Arizona's Enhanced I/M Program", presentation to National Research Committee to Review MOBILE Model, Irvine, CA, March 4.

Wenzel, Tom. 1999b. "Tracking Vehicles over Time in the Phoenix I/M Program", draft memo, April 16.

Wenzel, Tom. 1999c. "Evaluation of Arizona's Enhanced I/M Program", presentation at the 9th CRC On-Road Vehicle Emissions Workshop, San Diego, CA, April 21.

Figure 1. HC After Repair Adjustment Factors by MY
Proposed MOBILE6 vs. 1996-97 Arizona IM240

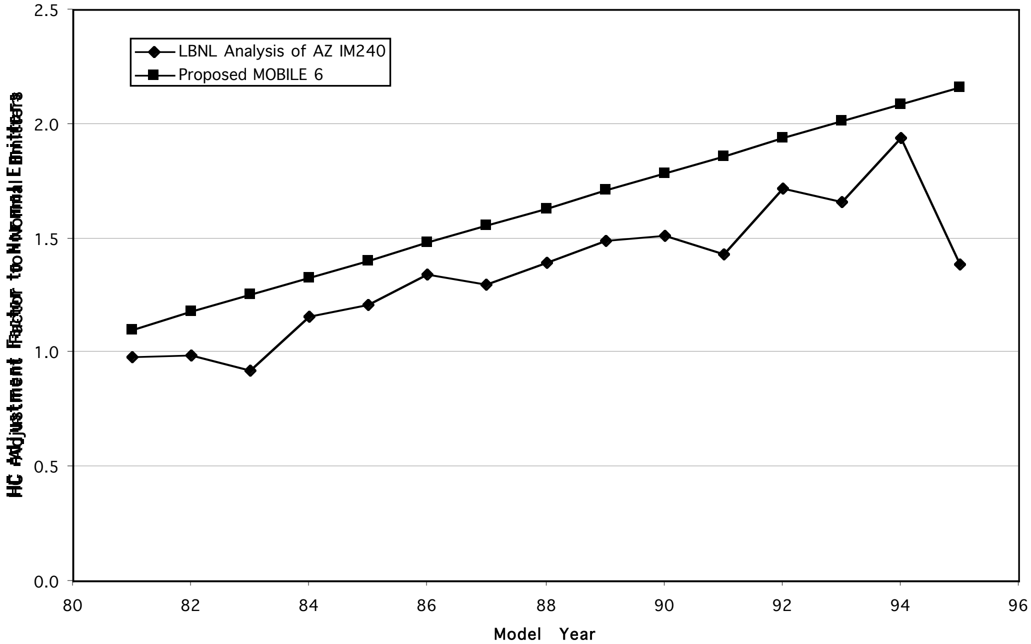


Figure 2. CO After Repair Adjustment Factors by MY
Proposed MOBILE6 vs. 1996-97 Arizona IM240

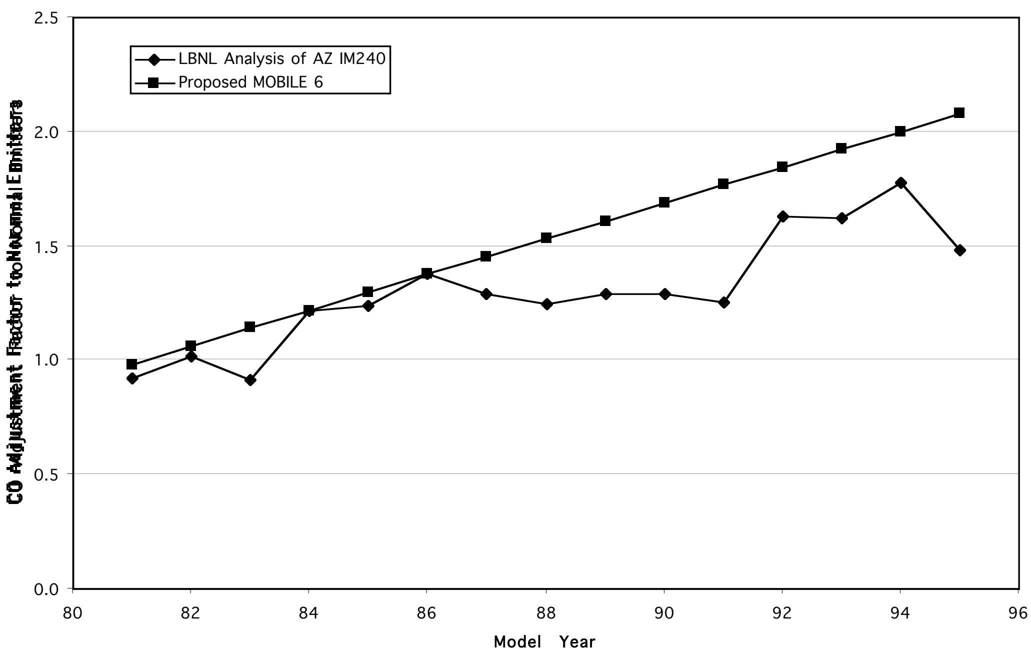


Figure 3. NOx After Repair Adjustment Factors by MY
Proposed MOBILE6 vs. 1996-97 Arizona IM240

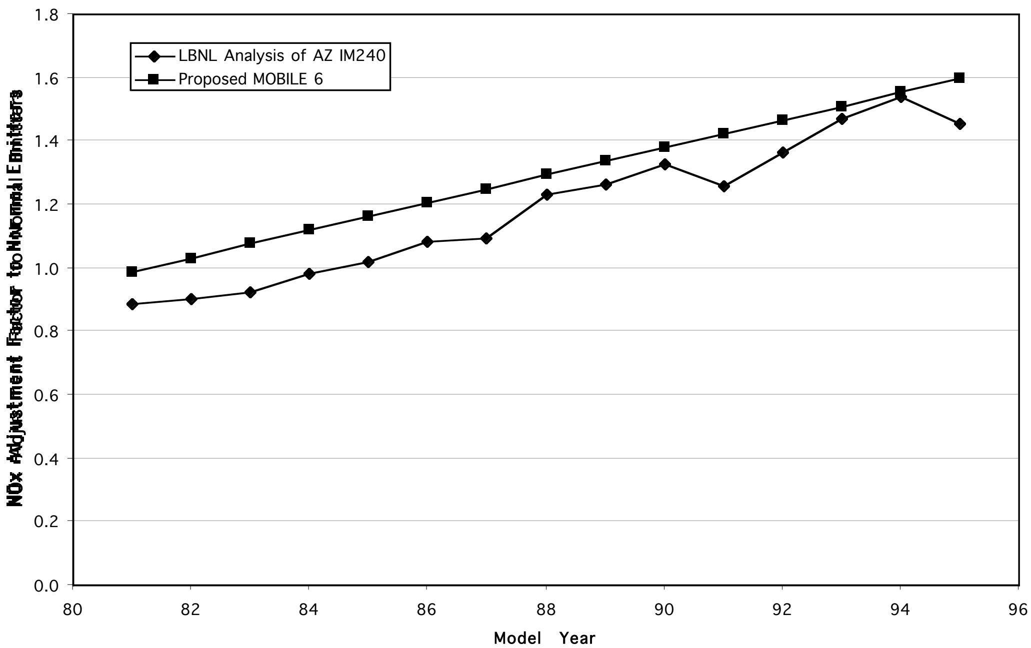


Figure 4. Proposed MOBILE6 HC Deterioration Rates, Cars

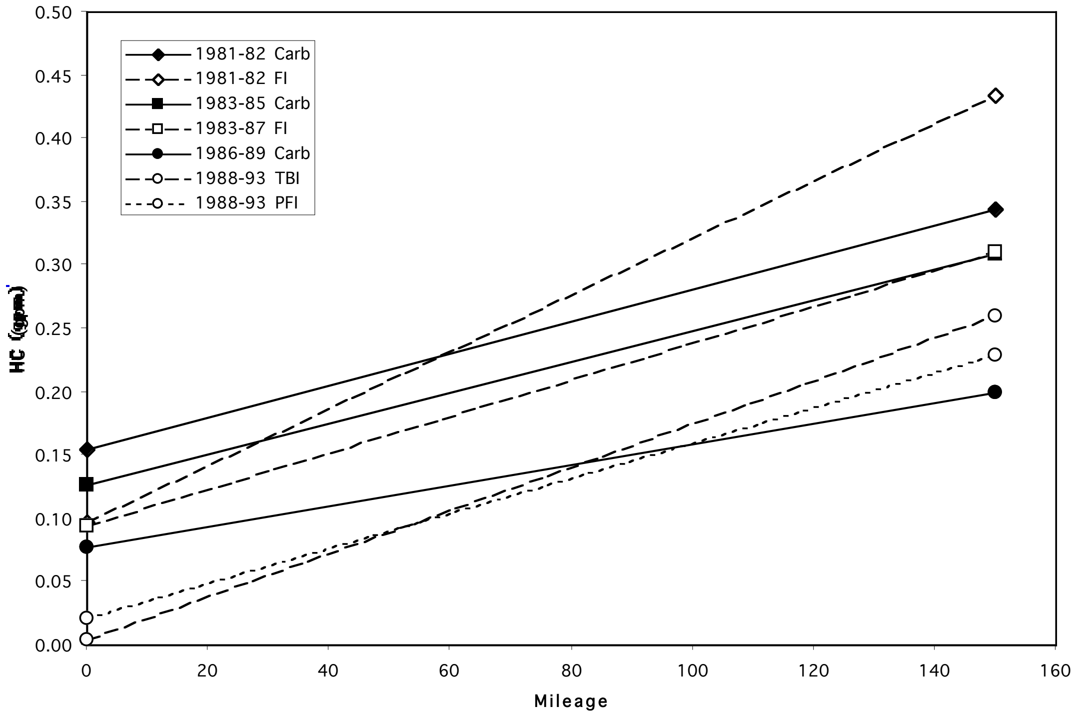


Figure 5. Proposed MOBILE6 CO Deterioration Rates, Cars

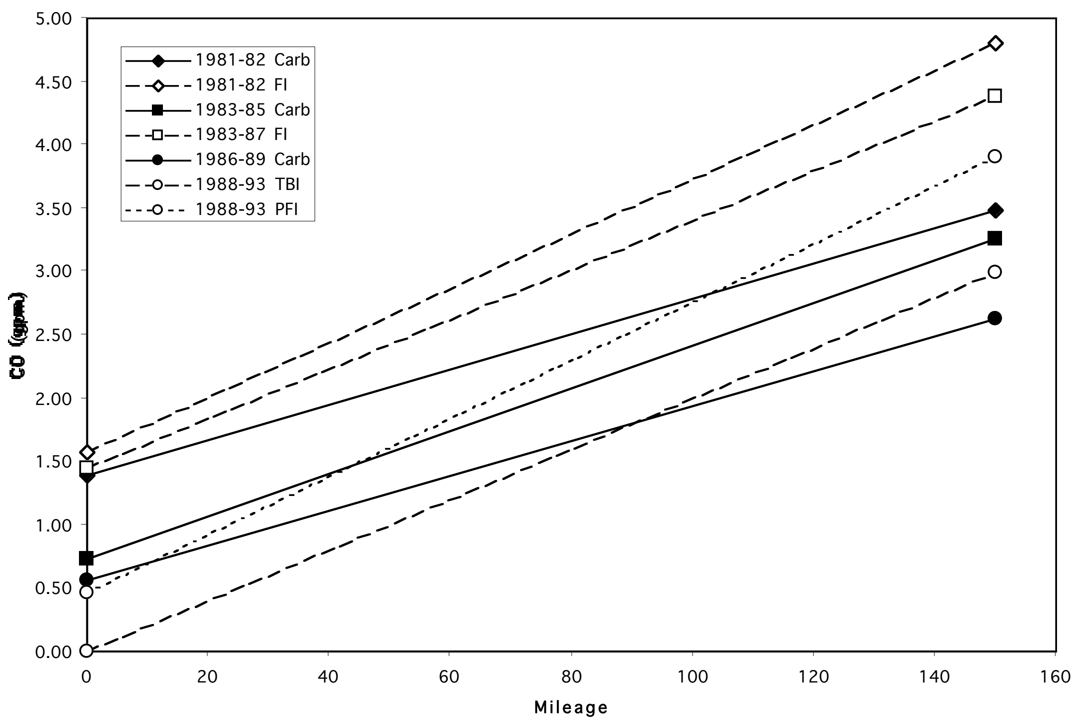


Figure 6. Proposed MOBILE6 NOx Deterioration Rates, Cars

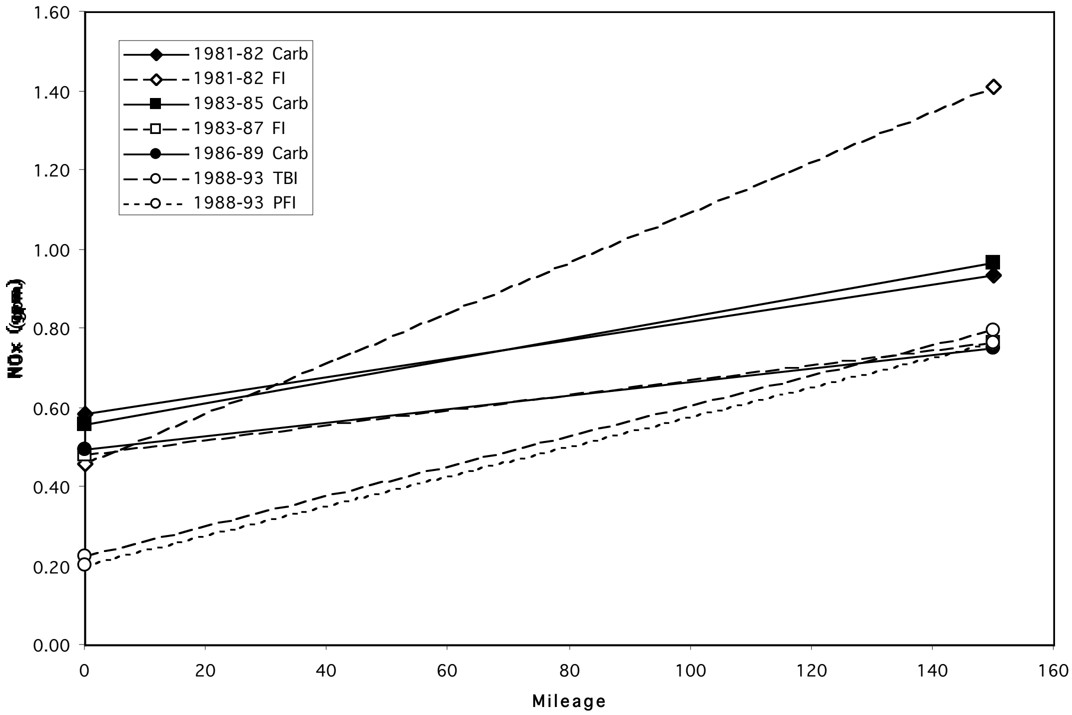


Figure 7. Proposed MOBILE6 HC Deterioration Rates, Light Trucks

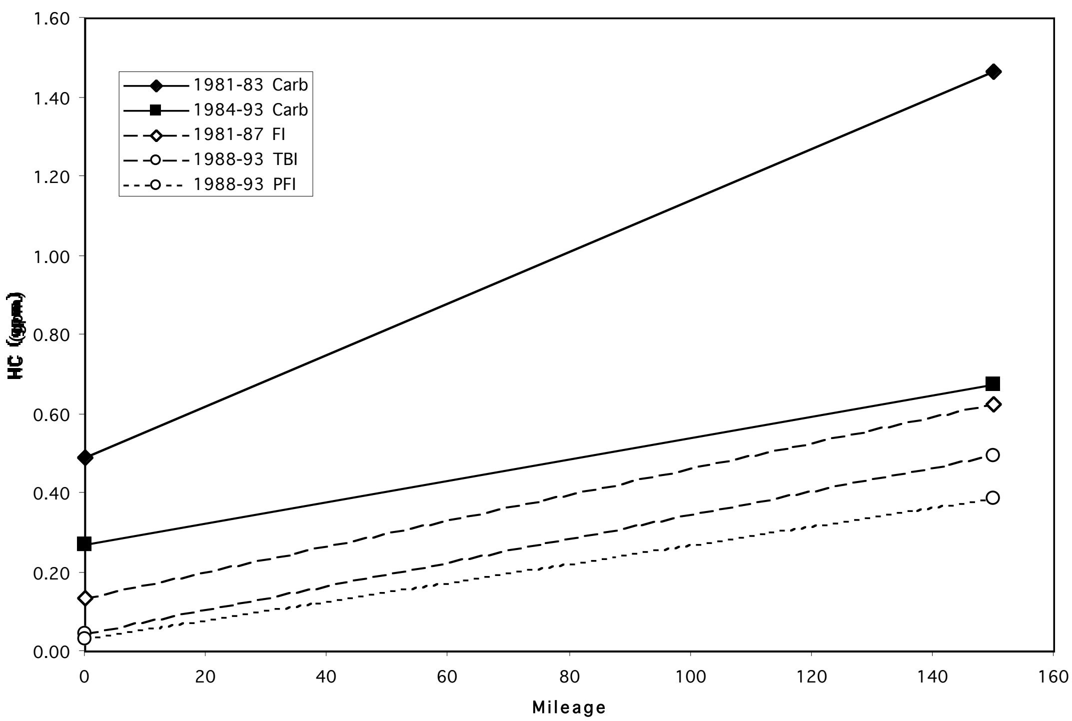


Figure 8. Proposed MOBILE6 CO Deterioration Rates, Light Trucks

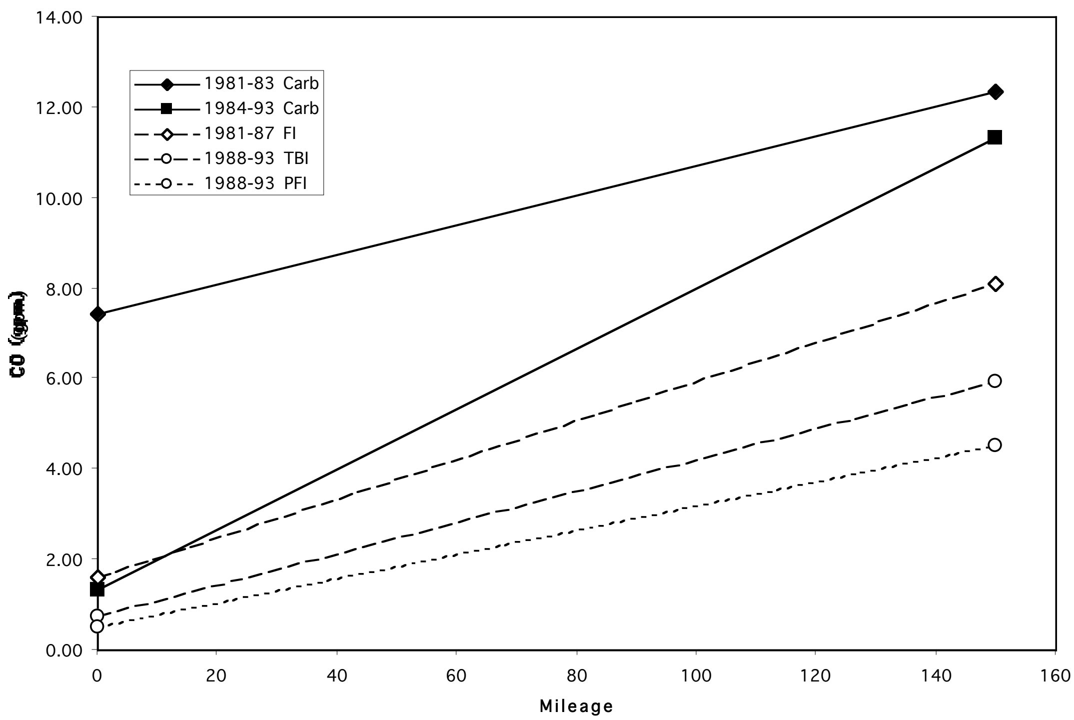


Figure 9. Proposed MOBILE6 NOx Deterioration Rates, Light Trucks

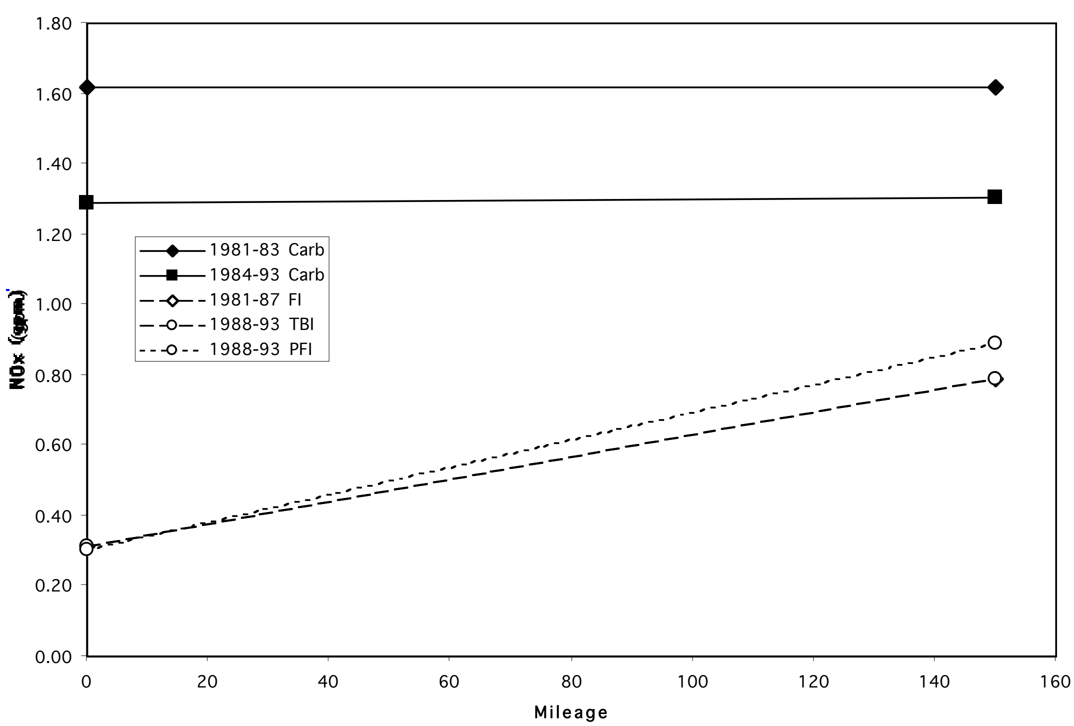
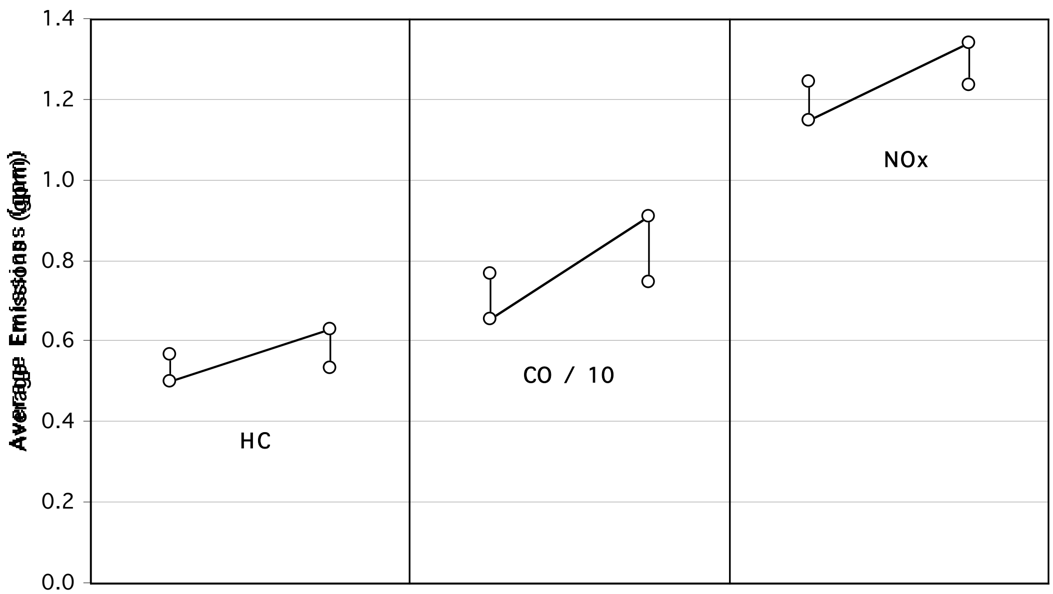


Figure 10. Annual I/M Credits Sawtooth
Passenger Cars, 1995 and 1997 Arizona IM240



APPENDIX O

APPENDIX P

LBNL-44159

Emission Reduction Potential from Repairing Gross Emitters

Tom Wenzel, Robert F. Sawyer

February 4, 1999

Energy Analysis Department
Environmental Energy Technologies Division
Lawrence Berkeley National Laboratory
University of California
Berkeley, CA 94720

This work was supported by the Office of Mobile Sources, Office of Air and Radiation, of the U.S. Environmental Protection Agency. Prepared for the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

Summary

This report describes our analysis of the emissions characteristics of vehicles that never complete I/M testing. These vehicles represent an emissions reduction potential that currently is lost by I/M programs. In this report we quantify these lost potential emissions reductions, and examine the effect of identifying these vehicles, using remote sensing, and repairing them. Our analysis consists of three steps:

1) compare the emissions of vehicles identified by remote sensing as "gross emitters" with those of "normal emitters";
2) calculate the total potential emission reductions lost by vehicles not completing I/M; and
3) estimate the fraction of these lost emission reduction that can be recovered by identifying these vehicles with remote sensing and repairing them.

The analysis is based on IM240 and remote sensing measurements of 412,000 model year 1981 and newer vehicles, measured in the Phoenix I/M area between January 1996 and June 1997.

Gross vs. Normal Emitters

We first examine the I/M emissions of gross emitters, to determine if they can be repaired down to the same emissions level as normal emitters. We examine 263,000 vehicles with remote sensing measurements prior to their initial I/M test. Of these vehicles, 27,400 (10%) are gross emitters, with at least one remote sensing measurement exceeding 4% CO or 500 ppm HC. We divided the vehicles into 4 groups, based on the result of their I/M testing:

1) vehicles that pass their initial IM240 test;
2) vehicles that fail their initial test, but pass a retest;
3) vehicles that fail their initial test, and fail subsequent testing; and
4) vehicles that fail their initial test and never receive a subsequent test.

Group 3 vehicles should include all vehicles that are waived from meeting IM240 standards, after having made repairs up to the repair cost limit. Technically, these vehicles should be excluded from our analysis, since they legally did not complete the I/M program. However, the number of waived vehicles is quite small, only 54 vehicles over the period studied (to be confirmed by AZ DEQ).

Table 1 shows the distribution of both gross emitting and normal emitting vehicles among these four groups, by vehicle type (passenger cars, light duty trucks less than 6,000 lbs GVW, and light duty trucks 6,000 to 8,000 lbs GVW).

Table 1. Number of Vehicles by Type, Emitter Type, and I/M Result

Type	Emitter Type	Group 1	Group 2	Group 3	Group 4	Total	Disn
Cars	Normal	133,608	8,225	2,379	1,894	146,106	90%
	Gross	10,637	2,819	1,533	1,149	16,138	10%
LDT1	Normal	66,220	2,941	514	471	70,146	89%
	Gross	7,150	1,235	388	326	9,099	11%
LDT2	Normal	18,886	873	114	120	19,993	90%
	Gross	1,566	408	100	68	2,142	10%

Table 2 shows the distribution of vehicles by I/M result. About 9% of cars with "normal" remote sensing emissions (that is, less than 4% CO and 500 ppm HC) fail initial I/M testing, and about two-thirds of those pass their final I/M test. In contrast, nearly 34% of gross emitter cars fail initial I/M testing, with only half of them passing out of the I/M program. There is a similar disparity in normal and gross emitter trucks, although the disparity is not as large as for cars.

Table 2. Distribution of Vehicles by Type, Emitter Type, and I/M Result

Type	Emitter Type	Group 1	Group 2	Group 3	Group 4	Total
Cars	Normal	91%	6%	2%	1%	100%
	Gross	66%	17%	9%	7%	100%
LDT1	Normal	94%	4%	1%	1%	100%
	Gross	79%	14%	4%	4%	100%
LDT2	Normal	94%	4%	1%	1%	100%
	Gross	73%	19%	5%	3%	100%

Figures 1 and 2 show average CO and HC emissions from initial and final I/M testing, for normal and gross emitting cars in each of the 4 groups. On average, gross emitters have higher emissions on their initial I/M test than normal emitters; among cars that fail initial I/M testing, gross emitters have initial emissions nearly twice that of normal emitters. The shaded columns indicate the level of emissions of the final I/M test for each group of vehicles (since vehicles in Groups 1 and 4 have only one I/M test, the "final" test is the same as the initial test). The difference between the clean and shaded columns for each group is the emissions reduction due to the I/M program. In general, cars that fail initial but pass final I/M testing (Group 2) see large reductions in emissions. (Not all of this reduction in emissions can be attributed to vehicle repairs. It is possible that a vehicle that was not properly warmed up, or preconditioned, prior to initial I/M testing was falsely failed, and passed a subsequent I/M test after sufficient preconditioning, with no repairs being made.) Failing cars that receive a second I/M test, but never pass out of the I/M program (Group 3), do show a small reduction in emissions.

Note that for both normal and gross emitters, cars that pass subsequent I/M testing (Group 2) show large average reductions in emissions. However, their emissions are not brought down to the levels of cars that pass their initial I/M test (Group 1). Although Group 2 gross emitters have substantially higher initial emissions than Group 2 normal emitters, their final emissions are only slightly higher. This suggests that gross emitters can be successfully repaired, or at least preconditioned to pass a second I/M test, bringing their emissions down to the level of normal emitters. Analysis of initial and final I/M emissions from light duty trucks 1 and 2 show results similar to those from cars.

Figure 3 shows the same data for NOx from cars. Normal emitters tend to have higher NOx emissions than gross emitters; this is because the definition of gross emitters is based on high CO or HC, and not NOx, remote sensing measurements. CO and HC emissions tend to correlate well, whereas NOx tends to be inversely correlated with CO and HC emissions. For example, nearly half of all CO failures also fail for HC, and 75% of HC failures also fail for CO, while only 25% of NOx failures fail for another pollutant as well. Therefore, it is not surprising that the "normal" emitter group, as defined by remote sensing measurements of CO and HC, includes vehicles with high NOx.

Part of the difference in post-I/M emissions of Group 1 and Group 2 vehicles may be due to different vehicle distributions by model year within each group. Figure 4 shows the distribution of cars by model year by I/M result (because their distributions are nearly identical, Groups 3 and 4 are combined). Group 1 vehicles tend to be much newer than the other vehicles, while Group 2 vehicles are slightly newer than Group 3 and 4 vehicles. Newer vehicles tend to have lower emissions than older vehicles, since they are built to meet tighter certification standards and have accumulated fewer miles. On the other hand, they are subject to tighter cutpoints in the I/M program. Figures 5 and 6 show average CO and HC emissions of gross emitting cars by I/M result and model year. Here we see that the emissions are reduced quite consistently for Group 2 cars across all model years. Final IM240 emissions of Group 2 cars are only slightly higher than emissions from Group 1 cars. (The differences between final IM240 emissions from Group 2 cars and Group 1 car emissions are larger in Figures 1 and 2 because Group 2 cars are substantially older than Group 1 cars). The figures also show the final IM240 emissions of Group 2 normal emitters, for comparison. Final IM240 emissions of Group 2 gross emitters are only slightly higher than those for Group 2 normal emitters.

Total Emission Reduction Potential

To quantify the total emission reduction potential of repairing Group 3 and 4 vehicles in the entire I/M fleet, we compared initial and final IM240 emissions of all 788,000 vehicles receiving their initial I/M test between January 1996 and June 1997. We assumed that the vehicles in Groups 3 and 4 would all be repaired, with their emissions reduced down to the post-I/M level of Group 2. This is an optimistic assumption, since the Group 3 and 4 vehicles have higher initial emissions than the Group 2 vehicles (Figures 5 and 6), and it may not be technically possible to repair their emissions down to the level of the Group 2 vehicles. We did this calculation by vehicle type and model year, and weighted the resulting emissions by annual vehicle miles traveled (VMT) for each vehicle type, using annual mileage data from Acurex 1997. We converted grams per year to (short) tons per day. Table 3 shows the emission reductions from repairing all 28,000 Group 3 and 4 vehicles in the Phoenix fleet. Initial IM240 emissions were 18.0 tpd HC and 261 tpd CO (these are reductions in tailpipe emissions; all calculations in this paper do not include evaporative HC emissions). Final IM240 emissions for the fleet were 15.6 tpd HC and 223 tpd CO, a 14% decrease attributable to the I/M program. If the Group 3 and 4 vehicles were identified and repaired, emissions would be reduced by an additional 11%, to 13.8 tpd HC and 196 tpd CO. Including the benefit of repairing the Group 3 and 4 vehicles nearly doubles the effectiveness of the I/M program.

Table 3. Emission Reductions in Tons per Day from Repairing All 28,000 Group 3 and 4 Vehicles (All Vehicles=788,000)

Pollutant	Initial IM240 tons/day	Final IM240 (Group 2 "Repaired") tons/day	Final IM240 (Group 2 "Repaired") tons/veh*	Groups 3 and 4 Repaired tons/day	Groups 3 and 4 Repaired tons/veh*	Percent Reduction: Final IM240	Percent Reduction: Groups 3&4 Repaired
Tailpipe HC	18.0	15.6		13.8			
abs. reduction		2.5	0.017	1.8	0.023	14%	11%
cum. reduction				4.3			24%
Tailpipe CO	261	223		196			
abs. reduction		38	0.260	27	0.341	14%	12%
cum. reduction				64			25%

* Divide by 365 days/year to convert tons/vehicle to tons/day/vehicle

Table 3 also shows the emission reductions in terms of total tons per vehicle repaired. Emissions of the 53,000 vehicles that initially failed and were repaired or otherwise passed their final I/M test were reduced on average by 0.017 tons per year HC and 0.26 tons per year CO. The emissions from the 28,000 Group 3 and 4 vehicles, if repaired, would be reduced by 0.023 tons per year HC and 0.341 tons per year CO. The per vehicle emissions reductions are larger for the Group 3 and 4 vehicles than the Group 2 vehicles because their initial emissions are higher, as shown in Figures 5 and 6.

This analysis is based on 788,000 vehicles with a single initial I/M test between January 1996 and June 1997. An additional 70,000 vehicles (9%), with multiple initial I/M tests or that failed visual I/M inspection only, were excluded from the analysis. We also excluded about 173,000 vehicles (20%) with either out of state, temporary, or no license plates. Finally, only about 75% of the vehicles participating in the biennial I/M program were tested during the 18-month period for which we have data. We make the assumption that the sample of vehicles excluded from our analysis is comparable to the vehicles we analyzed. Therefore, the ton per day values in Table 3 need to be adjusted to account for the vehicles not included in this analysis. Table 4 shows that the ton per day values should be increased by a factor of 1.64 to reflect total emissions of the Phoenix IM240 fleet.

Table 4. Calculation of Adjustment Factor to Encompass Entire IM240 Fleet

Jumber of Vehicles				Percent Additional Vehicles	Cumulative Adjustment Factor
	788,150				
lus	70,371	nultiple initial tests and visual failures =	858,521	9%	1.09
lus	173,494	out of state, temporary, or no license plate =	1,032,015	20%	1.31
lus	258,004	ests from late 1997 =	1,290,019	25%	1.64

How Much of Potential Reduction Can Be Achieved?

Table 3 shows that repairing vehicles that do not complete the I/M program can result in large emission reductions. However, what fraction of these vehicles can be identified and successfully repaired? For this analysis we return to the 263,000 vehicles that had remote sensing measurements <u>prior</u> to their initial IM240 test. We calculated the emissions reductions from repairing all Group 3 and 4 vehicles, and only those Group 3 and 4 vehicles that were identified by remote sensing as gross emitters.

As shown in Tables 5 and 6, initial IM240 emissions for the 263,000 were 6 tpd HC and 87 tpd CO. Final IM240 emissions for the fleet were 5.2 tpd HC and 75 tpd CO, a 13% decrease attributable to the I/M program. Table 5 shows the result of identifying and repairing the 3,600 gross emitting Group 3 and 4 vehicles: emissions would be reduced by an additional 5%, to 4.9 tpd HC and 71 tpd CO. Table 6 shows that by repairing all 9,000 of the Group 3 and 4 vehicles, the emissions reduction would be nearly twice as much (11%; 0.6 tpd HC and 8 tpd CO) as repairing just the gross emitters (5%; 0.3 tpd HC and 4 tpd CO). Note that the emissions reductions per vehicle for repairing all Group 2 vehicles, and all Group 3 and 4 vehicles, in Tables 5 and 6 are nearly identical to those for the entire sample in Table 3. The emissions reductions per vehicle for repairing the gross emitters (Table 5; 0.028 tons HC and 0.45 tons CO) are 20% to 30% higher than the reductions per vehicle for repairing all Group 3 and 4 vehicles (Table 6; 0.023 tons HC and 0.33 tons CO).

Table 5. Emission Reductions in Tons per Day from Repairing 3,600 Gross Emitting Group 3 and 4 Vehicles (All Vehicles=260,000)

	Initial	Final IM240		Groups 3 and 4		Percent Reduction	
	IM240	(Group 2 "Repaired")		Repaired		Final	Groups 3&4
Pollutant	tons/day	tons/day	tons/veh*	tons/day	tons/veh*	IM240	Repaired
Tailpipe HC	6.0	5.2		4.9			
abs. reduction		0.8	0.017	0.3	0.028	13%	5%
cum. reduction				1.0			17%
Tailpipe CO	87	75		71			
abs. reduction		11	0.248	4	0.450	13%	6%
cum. reduction				16			18%

* Divide by 365 days/year to convert tons/vehicle to tons/day/vehicle

Table 6. Emission Reductions in Tons per Day from Repairing All 9,000 Group 3 and 4 Vehicles (All Vehicles=260,000)

	Initial	Final IM240		Groups 3 and 4		Percent Reduction	
	IM240	(Group 2 "Repaired")		Repaired		Final	Groups 3&4
Pollutant	tons/day	tons/day	tons/veh*	tons/day	tons/veh*	IM240	Repaired
Tailpipe HC	6.0	5.2		4.6			
abs. reduction		0.8	0.017	0.6	0.023	13%	11%
cum. reduction				1.3			22%
Tailpipe CO	87	75		67			
abs. reduction		11	0.248	8	0.330	13%	11%
cum. reduction				19			22%

* Divide by 365 days/year to convert tons/vehicle to tons/day/vehicle

The distribution of vehicles, emissions, and emission reductions by vehicle type, emitter type, and I/M result are shown in Table 7. The table indicates that "normal" emitters as defined by remote sensing have as much emission reduction potential as gross emitters. For instance, 33% of the total potential HC emission reduction comes from gross emitting cars, whereas 35% comes from normal emitting cars. Table 7 also indicates that the majority of emissions reduction potential comes from cars (68% of HC, 73% of CO) as opposed to light duty trucks (32% of HC, 27% of CO).

In practice, not all gross emitters that do not complete I/M testing would be measured by remote sensing. Of all vehicles studied, 262,000 (64%) had remote sensing measurements after their final IM240, and 6,000 of these do not complete I/M testing. Reducing the emissions of these vehicles down to the final IM240 emissions of Group 2 vehicles results in an emissions reduction of 0.34 tpd HC and 5.0 tpd CO. Of the 262,000 vehicles with post-I/M remote sensing measurements, only 22,000 (8%) were gross emitters, and 2,196 of these did not pass their final IM240. Reducing the emissions of these vehicles results in emission reductions of 0.16 tpd HC and 2.6 tpd CO.

Limitations of This Analysis

The above analysis did not account for two important effects that would affect the emission reduction calculations. First, any repairs made on failing vehicles may not be durable. Analysis of three years of I/M data shows that 40% of vehicles that fail for any pollutant in 1995 fail again in 1997. The percentage of repeat failures ranges from 50% for MY81 vehicles to 10% for MY94 vehicles (Wenzel 1998). Because such a large fraction of vehicles that are supposedly repaired in the first round of I/M fail the second round, much of the emissions reduction quantified immediately after I/M will be lost over time. By ignoring the effect of repeat failures of the same vehicles, our analysis over-estimates the benefits from repairing Group 3 and 4 vehicles.

On the other hand, it is possible that vehicles that never pass I/M are removed from the I/M area, either through resale out of the area or scrappage. In an earlier analysis, we looked at the populations of vehicles seen by remote sensing in multiple periods after each vehicle's final I/M test (Wenzel 1998). The fraction of Group 2 and Group 3/4 vehicles seen by remote sensing decreases the further one gets from the final I/M test. By 6 months after the final I/M test, the Group 3 and 4 portion of the fleet is reduced by 40%; only one-third of the Group 3 and 4 vehicles are still driven in the I/M area over 15 months after their final I/M test. Because Group 3 and 4 vehicles tend to drop out of the fleet at a greater rate than other vehicles, fewer of these vehicles are available for repair to reduce emissions, and our analysis over-estimates the effect of repairing gross emitters (however, the removal of these vehicles from the I/M area, perhaps as a result of the I/M program, does represent an emission reduction typically not quantified in current evaluations of I/M programs). One would expect that the oldest vehicles are the ones that are being "retired" from the I/M area. However, our analysis indicates that the model year distribution of remote sensing readings 15 months after I/M testing is nearly identical to the distribution of readings immediately after I/M testing (Wenzel 1998).

This analysis was restricted to model year 1981 and newer vehicles tested over 18 months of a biennial I/M program. Three groups of vehicles were not included in the analysis: a) model year 1980 and older vehicles, that receive an idle test rather than an IM240; b) vehicles not scheduled for I/M testing until the second half of 1997; and c) vehicles not participating in the I/M program (either legally registered outside of the I/M area, or not registered). We attempted to account for vehicles in group b), by developing an adjustment factor to increase the ton per day emissions values we calculated. However, the other two groups of unaccounted for would increase the baseline emissions inventory, and could affect the calculated emissions reductions from repairing vehicles that do not complete I/M.

The remote sensing data used in this analysis provides only the license plate of the measured vehicle; to obtain vehicle information, remote sensing records must be matched with IM240 records, by license. Therefore there is no information regarding vehicles measured by remote sensing that do not appear in the IM240 database (groups a and c, described above, as well as out of state vehicles that become registered in the area. In addition, if vehicle owners switch license plates between remote sensing measurement and I/M test, remote sensing readings will be assigned to the wrong vehicle and I/M test result. This should not be a major problem, since in Arizona license plates stay with vehicles, rather than drivers, when a vehicle is sold (in states like Colorado license plates stay with drivers, not vehicles). Remote sensing data could be made more accurate by regularly matching license plates with registration information as the data are collected.

Finally, our analysis only examines the maximum emission reduction potential from the vehicles that do not complete the I/M program. The analysis does not consider if these emissions reductions can actually be achieved through vehicle repair, or whether it would be cost-effective to do so.

Issues for Further Analysis

There are several ways this analysis could be improved to evaluate the effect of repairing vehicles that do not complete I/M. The analysis used a definition of gross emitter as a that exceeded CO or HC remote sensing cutpoints at least once. Further analysis could require at least 2 remote sensing exceedances per vehicle. There are 356,000 vehicles with at least 2 remote sensing readings either before or after I/M testing; 13,000 of these vehicles exceed the gross emitter cutpoints at least twice. In addition, we could use CO cutpoints only in defining gross emitters; earlier research indicates that there are some potential problems with the HC remote sensing data from Phoenix, including negative readings rounded to zero and many deceleration sites resulting in high HC readings (Wenzel, 1998). Another possibility is to use remote sensing cutpoints that vary by model year, so that more newer vehicles are included as gross emitters eligible for repair.

Since remote sensing does not, and indeed cannot, identify all of the vehicles not completing I/M, another approach would be to subsidize repair on the highest emitters, as measured by IM240 during I/M testing. Higher IM240 cutpoints, either constant or varying by model year, could be established; any vehicle exceeding the cutpoints would be eligible for repair by better trained mechanics. Such an approach would ensure that the highest emitters are identified, and are repaired while they are still participating in the I/M program.

Finally, a more detailed analysis of repair effectiveness could be performed by comparing IM240 emissions of the same vehicle over multiple I/M cycles. Such an analysis would separate vehicles into normal and gross emitters, on the basis of either remote sensing or IM240 emissions, and compare the long-term repair effectiveness of each group.

Conclusions

This analysis indicates that nearly half of the potential reduction in HC and CO emissions from the Phoenix I/M program is lost by not fully repairing vehicles that do not complete I/M testing. Most of the lost emission reductions comes from cars rather than light duty trucks. Only half of the lost emission reductions can be attributed to vehicles with high remote sensing readings; the other half of the lost reductions comes from vehicles with normal remote sensing readings. Only 64% of the vehicles that do not complete I/M were measured by remote sensors after their I/M test, and only 8% of these vehicles were gross emitters. Because normal emitters account for half of the lost emission reductions, it may not be efficient to use remote sensing measurements to identify vehicles eligible for repair assistance.

References

Acurex Environmental Corporation. 1997. *Update of Fleet Characterization Data for Use in MOBILE6*. May 16.

Wenzel, Tom. 1998. “Evaluation of Arizona’s Enhanced I/M Program.” Presentation at the Clean Air Conference. Breckenridge, Colorado. September 18.

APPENDIX Q

Using Program Test Result Data to Evaluate the Phoenix I/M Program

Report to the Arizona Department of Environmental Quality

December 9, 1999

Tom Wenzel
Lawrence Berkeley National Laboratory
1 Cyclotron Road
Room 90-4000
Berkeley, CA 94720

Executive Summary

This report uses emissions test result data from 1997 to evaluate the effectiveness of the enhanced I/M program in reducing vehicle tailpipe emissions in Phoenix, Arizona. The analysis is based on a comparison of initial and final test results for individual vehicles that received their initial I/M test in 1997. Two types of tests are performed on vehicles subject to I/M testing in Phoenix; the idle and loaded idle test is required of 1980 and older vehicles, while 1981 and newer vehicles must take the IM240 test. Significant differences between the two types of test require that the emissions of the two fleets be analyzed separately.

Arizona allows vehicles to fast pass or fast fail the IM240 test; in order to compare emissions of vehicles tested over different portions of the IM240, we must convert these “short test” results to full IM240 test equivalents. A relatively simple method to make this conversion is used; a comparison of this method with other more detailed methods indicates that all methods tend to underestimate full IM240 emissions using fast pass/fast fail emissions results. The analysis does not consider the effect of the I/M program on reducing evaporative HC emissions.

Comparison of initial and final IM240 tests indicates that the program is reducing the average per vehicle emissions by 16% for HC, 17% for CO, and 7% for NOx, for the entire vehicle fleet. After weighting per vehicle emissions by estimated annual miles traveled, the fleetwide emissions reductions are 2.3 tons per day (14% reduction) for HC, 34 tons per day (15% reduction for CO), and 2.3 tons per day (7% reduction) for NOx. CO and NOx reductions appear to be substantially larger for cars than for light duty trucks. Per vehicle emissions of the loaded idle fleet are reduced by 15% for HC and 23% for CO.

About 11% of all vehicles fail their initial IM240 emissions test; the failure rate is slightly higher for passenger cars (12%) than for light duty trucks (8%). The initial failure rate for the loaded idle test is 37%. Of the vehicles that fail their initial test, only 70% received a final passing test through March 1998; 30% did not receive a final passing test through March 1998. Because waivered vehicles are not identified in the data, the actual percentage of No Final Pass vehicles is likely to be closer to 26%. The percentage of No Final Pass cars is greater than the percentage of No Final Pass trucks.

The percent reductions in loaded idle emissions for Final Pass vehicles tend to increase by model year, with larger reductions for newer vehicles. There is a large increase in percent reduction for model year 1974 through 1980 vehicles, presumably due to stricter cutpoints applied to those vehicles. The percentage reductions of IM240 Final Pass vehicles from model years 1981 through 1993 are fairly constant by model year. HC and CO emission reduction percentages tend to increase after model year 1993.

We use a relatively crude method to estimate total emissions and emission reductions in tons per day for the loaded idle fleet, in order to estimate the tonnage reductions for the entire Phoenix I/M program. We estimate that the program reduces the emissions of the fleet reporting for I/M by 3.0 tons per day for HC, 38 tons per day for CO, and 2.6 tons per day for NOx. The majority of the estimated emissions reductions comes from the IM240 fleet: 76% for HC, and 88% for CO and NOx. The estimated percent reduction in total emssions is 15% for HC, 13% for CO, and 7% for NOx.

The estimated effectiveness of the I/M program depends on whether the No Final Pass vehicles have been permanently removed from the I/M area, or if they continue to be driven in the I/M area. The effectiveness of the program on the IM240 fleet nearly doubles if one assumes that all IM240 No Final Pass vehicles have been permanently removed from the area. Analysis of 1995 IM240 test data and remote sensing data indicate that about half of the No Final Pass vehicles continue to be driven in the I/M area. If this information is correct for vehicles tested in 1997, the 1997 I/M program resulted in a 22% reduction in HC and CO, and a 9% reduction in NOx from the IM240 fleet. These percentage reductions are equivalent to 3.0 tons per day for HC and NOx, and 48 tons per day for CO.

Analysis of a single year of I/M program test data can only provide a partial understanding of the program's effectiveness in reducing emissions. Tracking of individual vehicles over several I/M cycles can reveal important information on long-term effectiveness of vehicle repair, and changes in the fleet reporting for I/M testing. In addition, an independent source of on-road emissions tests, such as from a remote sensing measurement program, can provide additional information on repair effectiveness, the effect of pre-test repairs on emissions, and the number and emissions of vehicles avoiding the I/M program.

1. Introduction

This report uses emissions test result data from 1997 to evaluate the effectiveness of the enhanced I/M program in reducing vehicle tailpipe emissions in Phoenix, Arizona. Effectiveness is measured in terms of both percent and absolute tons of emissions reduced. The analysis is based on a comparison of initial and final test results for individual vehicles, on either the IM240 or the loaded idle test, depending on the age of the vehicle.

Model year 1981 and newer vehicles with two-wheel drive are subject to IM240 dynamometer testing in the Phoenix I/M program. Model year 1967 to 1980 vehicles registered in the Phoenix area are subject to an idle and a loaded idle I/M test. Both idle test emissions are reported as pollutant concentrations in the exhaust (percent for CO, parts per million for HC), which are not directly comparable to the mass emissions (grams per mile) reported from IM240 tests. In addition, NOx emissions are not measured under the idle tests. Because of these differences between the two tests, we analyze the fleet of vehicles subject to each type of test separately. For the pre-1981 vehicles we use emissions from the loaded idle test, since this test is somewhat more similar to the IM240 test than the conventional idle test. Because they cannot be driven on the dynamometers used for IM240 or loaded idle testing, all-wheel drive vehicles of all model years are subject to an idle test only. 13,000 such vehicles registered in the Phoenix area were tested in 1997; nearly 90% of these vehicles are 1981 and newer. We exclude all of these all-wheel drive vehicles from our analysis.

There is another important difference between the test results for loaded idle and IM240 tests. Vehicles subject to the IM240 test are classified as either passenger cars, light duty trucks less than 6,000 pounds, or light duty trucks between 6,000 pounds and 8,500 pounds. However, vehicles subject to loaded idle testing are classified as either: 1) less than 6,000 pounds and 4 or fewer cylinders; 2) less than 6,000 pounds and more than 4 cylinders; or 3) between 6,000 pounds and 8,500 pounds. Therefore, comparison of the IM240 and loaded idle fleets by vehicle type requires that the first two classifications (passenger cars and light duty trucks under 6,000 pounds) be merged into a single group.

The next section describes the process used to convert IM240 short test emission results to full IM240 equivalent emissions levels. Section 3 presents estimates of program effectiveness by vehicle type/class, for each of the IM240 and loaded idle vehicle fleets; Section 4 presents program effectiveness for each fleet by I/M test result. In Section 5 we combine the data from the analysis of the two independent fleets to derive estimates of program effectiveness on all vehicles reporting for I/M testing. Section 6 discusses how vehicles that never complete I/M testing affect the evaluation of program effectiveness. Other issues critical to accurate evaluation of I/M programs, but not specifically addressed here, are discussed in Section 7. Section 8 summarizes our results and provides some conclusions.

2. IM240 Short Test Conversion

This analysis is based on all initial IM240 tests of vehicles performed in 1997, with the exceptions described below. Arizona allows vehicles to either "fast pass" the IM240 after only 31 seconds of testing, or "fast fail" the test after 94 seconds of testing. Therefore, virtually all vehicles are either passed or failed before they complete the full 240 seconds of the IM240 test. To compare emissions of vehicles tested over different portions of the IM240, we must convert these "short test" results to full IM240 test equivalents.

We used a rather simple method to make this conversion; we obtained from EPA second-by-second full IM240 test results on 4,000 vehicles conducted by Automotive Testing Laboratories (ATL) in Arizona in 1992. Figure 1 shows the speed time trace of the IM240 (right scale), and the average gram per mile emissions of the ATL test fleet at each second of the test (left scale). For each second of the test, cumulative grams are divided by cumulative miles for each vehicle, and the results are averaged over the fleet. The highest average gram per mile values occur at second 30, and decrease as the test continues. The hardest acceleration in the IM240 occurs just before second 160; this acceleration causes the cumulative average gram per mile values for CO and NOx to increase slightly.

We then calculated the ratio of the emissions at each second to the emissions for the full IM240, for each pollutant for each vehicle. Figure 2 shows the ratios averaged over all vehicles, for each pollutant; we use these average ratios as adjustment factors to convert short test results to full test equivalent emissions. The adjustment factors are quite large for vehicles passed immediately after 30 seconds; for example, for these vehicles we divided measured HC gram per mile values by 3.4 to obtain full-IM240 equivalent HC emissions. Each of the adjustment factor curves reaches 1 at second 240, indicating that no adjustments were made to vehicles driven the full 240 seconds of the test.

Figure 1. Average gpm Emissions at Each Second of IM240
ATL Arizona Data

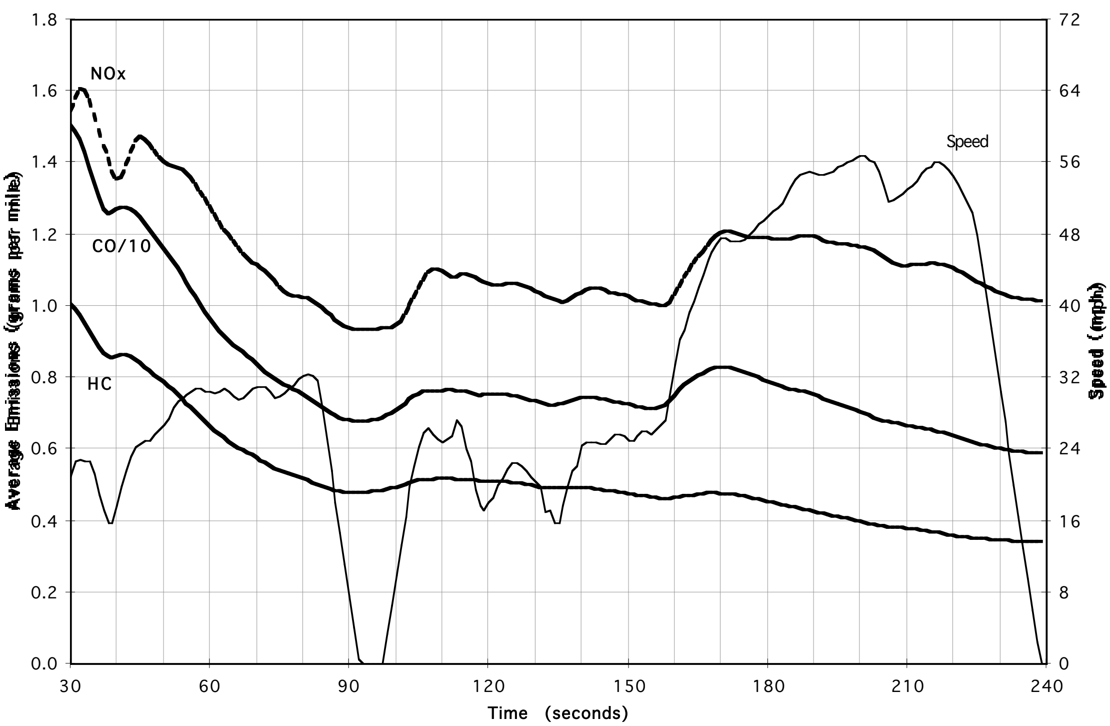
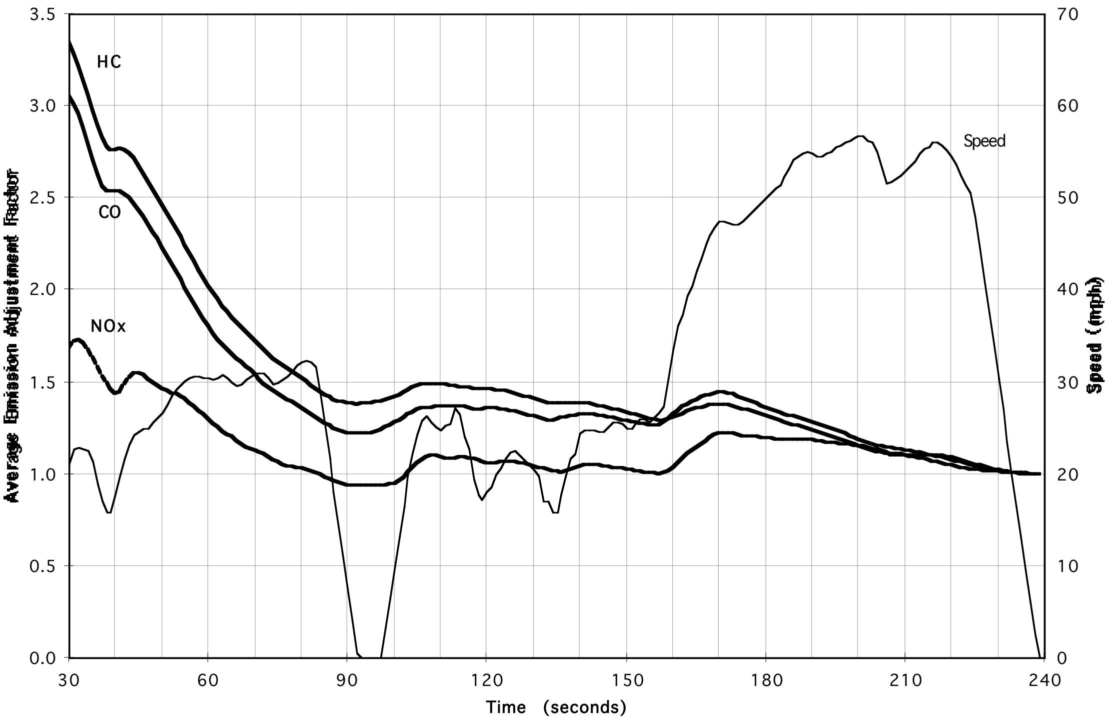


Figure 2. Average Emission Adjustment Factor for Each Second,
ATL Arizona Data



Our method involves dividing measured emissions at a given second by a conversion factor, based only on the second of testing. Others have developed different, more involved methods for converting short test emissions, using other variables such as vehicle type and age. We have compared several different methods for converting short test emissions to full IM240 equivalents (the comparison is included as Appendix A of this report). This comparison found that all of the methods tend to underestimate full IM240 emissions of fast pass vehicles. One reason for the underestimation is that a small number of vehicles (one or two percent) are improperly fast passed; if allowed to complete the full IM240 test, their emissions would exceed the full IM240 cutpoints. In general, all of the conversion methods are more accurate for vehicles tested over longer segments of the IM240 test. Since Arizona does not fail high emitters until at least second 94 of the IM240, we believe the adjustment is more accurate for the failing vehicles than vehicles passed immediately after second 30.

3. Initial Program Effectiveness by Vehicle Type/Class

To estimate initial effectiveness of the Phoenix program, we compared the initial and last test of each vehicle with an initial test in 1997. To do this we first matched all vehicle tests by vehicle identification number (VIN). For vehicles with subsequent retests, we took the last retest through March 1998 as the final test of the vehicle. For vehicles that passed their initial test, and vehicles that failed their initial test but did not receive a retest, we assumed that their emissions were equivalent to those measured during their initial test. We excluded from our analysis 4,000 IM240 tests with invalid VINs[1] (less than 1% of all tests) and 18,000 vehicles (or 2.5% of all unique vehicles) with subsequent tests coded as initial tests[2,3]. Excluding these vehicles from our

---

1. The VIN has a check digit that can be used to determine if the combination of numerals and characters in the VIN are valid. Less than one percent of the vehicles had an invalid VIN.
2. There are several reasons why a vehicle may have multiple initial tests within a two-year period: vehicles for sale by dealers that are not fleet-licensed must be tested every 90 days; subsequent tests of vehicles that were not passed within 5 months of the initial test are coded as initial tests; some repeat initial tests are for research purposes only; a

analysis has little effect on average emissions per vehicle, but has a larger effect on absolute tons of emissions.

Table 1 shows the average initial and final emissions, in adjusted grams per mile, of passenger cars, light duty trucks less than 6,000 pounds GVW (LDT1), and light duty trucks between 6,000 and 8,500 pounds GVW (LDT2) tested on the IM240 in 1997. Table 2 shows the same data for the vehicles subject to the idle test. The table also shows the percentage emissions reduction for each vehicle type, and for the fleet as a whole, as measured by comparing the initial test with the final test of each vehicle. The tables indicates that the Phoenix I/M program is reducing emissions of the IM240 fleet by 16% for HC, 17% for CO, and 7% for NOx; the loaded idle emissions of the idle fleet are reduced by 15% for HC and 23% for CO.[4] The percentage reduction in IM240 CO and NOx, and the percentage reduction in loaded idle CO, appear to be substantially larger for cars than for light duty trucks. (This analysis does not consider evaporative HC emissions, and therefore understates the program's effectiveness in reducing total HC).

Table 1. Average Emissions and Percent Reduction, IM240 Fleet, Unweighted by Annual VMT

Type	Number	Unweighted Average Emissions per Vehicle (adjusted grams per mile)						Percent Reduction		
		HC		CO		NOx				
		Initial	Final	Initial	Final	Initial	Final	HC	CO	NOx
Cars	431,098	0.62	0.52	8.98	7.23	1.26	1.15	16.6%	19.5%	8.6%
LDT1	185,888	0.85	0.73	11.97	10.41	1.62	1.52	14.2%	13.0%	6.1%
LDT2	53,789	1.06	0.88	14.50	12.44	2.18	2.08	16.8%	14.2%	4.5%
All	670,775	0.72	0.61	10.25	8.53	1.43	1.33	15.8%	16.8%	7.3%

Table 2. Average Emissions and Percent Reduction, Loaded Idle Fleet, Unweighted by Annual VMT

Type	Number	Unweighted Average Emissions per Vehicle (emissions concentration)					
		HC (ppm)		CO (%)		Percent Reduction	
		Initial	Final	Initial	Final	HC	CO
Class 3	15,774	145	121	1.53	1.13	16.1%	26.0%
Class 4	66,573	113	95	1.17	0.90	15.6%	23.0%
Class 5	23,653	113	97	1.28	1.03	14.1%	19.5%
All	106,000	118	100	1.25	0.96	15.4%	22.7%

Tables 1 and 2 show the average emissions per vehicle; however, for inventory purposes, the per vehicle emissions reductions have to be weighted by the average number of annual miles driven by different types and ages of vehicles. Table 3 shows the average IM240 emissions from Table 1 in terms of tons per day, using EPA's latest estimates of annual vehicle miles traveled (VMT) by vehicle type and age (Acurex, 1997). Table 3 indicates slightly lower emissions reductions than Table 1. The absolute tons per day values in Table 3 may not be directly comparable to estimates of the Arizona mobile source emissions inventory, since the method to adjust the emissions of fast pass vehicles tends to underestimate full IM240 emissions of the majority of

small number of audit vehicles are covertly run through the system periodically; and a prospective buyer may voluntarily test a vehicle prior to purchase (personal communication with Frank Cox, Arizona Department of Environmental Quality).

3. We did not exclude any loaded idle tests because of invalid VINs, because the VIN was not standardized across all vehicle manufacturers until the 1981 model year. We did exclude 40,000 vehicles subject to the loaded idle test with multiple initial tests.

4. Idle emissions of the idle fleet are reduced by 25% for HC and 30% for CO.

vehicles, as described above. In addition, vehicles with invalid VINs and with multiple initial tests have been excluded, as described above. Again, percentage CO and NOx reductions appear to be substantially larger for cars than for light duty trucks.

Table 3. Total Emissions and Percent Reduction, Weighted by Annual VMT

Type	Number	Total Emissions (Tons per Day)						Percent Reduction		
		HC		CO		NOx				
		Initial	Final	Initial	Final	Initial	Final	HC	CO	NOx
Cars	431,096	8.7	7.4	125.7	103.1	18.8	17.3	15.3%	18.0%	7.6%
LDT1	185,885	5.1	4.4	72.9	64.8	10.6	9.9	13.0%	11.2%	6.0%
LDT2	53,788	2.0	1.7	26.9	23.7	4.8	4.6	14.3%	11.9%	3.7%
All	670,769	15.8	13.5	225.6	191.7	34.1	31.9	14.4%	15.0%	6.6%
Reduction			*2.3*		*34.0*		*2.3*			

Note: Absolute tons of emissions may not be comparable to official emissions inventories, due to conversion of fast pass/fast fail emissions to full IM240 emissions and exclusion of vehicles with invalid VINs, multiple initial tests, or that do not report for I/M testing.

As discussed above, we cannot calculate tons per day of the vehicles subject to the loaded idle test. In addition, we cannot calculate average idle emissions weighted by annual vehicle miles traveled, as the VMT assumptions we use vary by vehicle type as well as model year, and the loaded idle data are not classified by the same vehicle types. We return to this issue in Section 6.

4. Initial Program Effectiveness by I/M Result

As discussed above, we determined the final I/M result of each vehicle initially tested in 1997. We grouped vehicles into four groups, based on their first and last emissions test[5]:

1) vehicles that passed their initial test ("Initial Pass");
2) vehicles that failed their initial test, but passed a subsequent retest ("Final Pass")[6];
3) vehicles that failed their initial test and failed a subsequent retest ("No Final Pass"); and
4) vehicles that failed their initial test and had no retest ("No Second Test").

We frequently treat groups 3 and 4 as a single group, No Final Pass vehicles.

Tables 4 and 5 show the number and distribution of vehicles by vehicle type/class and I/M result. No Final Pass and No Second Test vehicles are shown separately, and grouped together and shown in italics. Table 4 indicates that about 11% of all vehicles fail their initial IM240 emissions test; the failure rate is slightly higher for passenger cars (12%) than for light duty trucks (8%). Table 5 shows that nearly three times as many vehicles fail their initial idle or loaded idle test (37%); again, the idle failure rate is higher for Class 5 vehicles (LDT2; 36%) than Class 3 vehicles (cars and LDT1 with 4 or fewer cylinders; 44%). Of the vehicles that fail their initial IM240 test, only 70% received a final passing test in 1997; 30% did not receive a final passing test in 1997. The percentage of IM240 No Final Pass cars is greater than the percentage of No Final Pass trucks (33% for cars, 23% for LDT1, 21% for LDT2). The overall No Final Pass rate for vehicles subject to loaded idle testing is similar to that for IM240 vehicles, with the No Final Pass rate decreasing as the class increases (38% for Class 3, 28% for Class 4, and 24% for Class5).

5. About 4% of all IM240 vehicles, and 20% of all loaded idle vehicles, passed their initial emissions test but failed either a functional or visual test; these vehicles are excluded from our analysis by I/M result.
6. Presumably emissions controls malfunctions are identified and repaired for most of these vehicles; however, it is possible that a number of these vehicles pass a retest without any permanent repairs being made.

Table 4. Number of IM240 Vehicles by Type and I/M Result*

Type	I/M Result	Number	Percent of Total	Percent of Initial Fails
Cars	1) Initial Pass	365,983	87.8%	
	2) Final Pass	33,912	8.1%	66.7%
	3) No Final Pass	9,348	2.2%	18.4%
	4) No Second Test	7,575	1.8%	14.9%
	Subtotal 3 and 4	*16,923*	*4.1%*	*33.3%*
	All Cars	416,818	100.0%	100.0%
LDT1	1) Initial Pass	161,450	91.8%	
	2) Final Pass	11,112	6.3%	77.2%
	3) No Final Pass	1,829	1.0%	12.7%
	4) No Second Test	1,448	0.8%	10.1%
	Subtotal 3 and 4	*3,277*	*1.9%*	*22.8%*
	All LDT1	175,839	100.0%	100.0%
LDT2	1) Initial Pass	45,694	91.6%	
	2) Final Pass	3,283	6.6%	78.8%
	3) No Final Pass	446	0.9%	10.7%
	4) No Second Test	438	0.9%	10.5%
	Subtotal 3 and 4	*884*	*1.8%*	*21.2%*
	All LDT2	49,861	100.0%	100.0%
All Vehicles	1) Initial Pass	573,127	89.2%	
	2) Final Pass	48,307	7.5%	69.6%
	3) No Final Pass	11,623	1.8%	16.8%
	4) No Second Test	9,461	1.5%	13.6%
	Subtotal 3 and 4	*21,084*	*3.3%*	*30.4%*
	Total	642,518	100.0%	100.0%

*Excludes 4% of vehicles that pass initial emissions test but fail initial visual or functional test.

Table 5. Number of Loaded Idle Vehicles by Class and I/M Result*

Class	I/M Result	Number	Percent of Total	Percent of Initial Fails
Class 3 (Cars and LDT1 with 4 or fewer cylinders)	1) Initial Pass	8286	55.5%	
	2) Final Pass	4122	27.6%	61.9%
	3) No Final Pass	1465	9.8%	22.0%
	4) No Second Test	1069	7.2%	16.1%
	Subtotal 3 and 4	*2,534*	*17.0%*	*38.1%*
	All Class 3	14,942	100.0%	100.0%
Class 4 (Cars and LDT1 with more than 4 cylinders	1) Initial Pass	39579	64.7%	
	2) Final Pass	15508	25.3%	71.8%
	3) No Final Pass	3405	5.6%	15.8%
	4) No Second Test	2699	4.4%	12.5%
	Subtotal 3 and 4	*6,104*	*10.0%*	*28.2%*
	All Class 4	61,191	100.0%	100.0%
Class 5 (LDT2)	1) Initial Pass	13668	63.9%	
	2) Final Pass	5889	27.5%	76.2%
	3) No Final Pass	1005	4.7%	13.0%
	4) No Second Test	839	3.9%	10.8%
	Subtotal 3 and 4	*1,844*	*8.6%*	*23.8%*
	All Class 5	21,401	100.0%	100.0%
All Vehicles	1) Initial Pass	61,533	63.1%	
	2) Final Pass	25,519	26.2%	70.9%
	3) No Final Pass	5,875	6.0%	16.3%
	4) No Second Test	4,607	4.7%	12.8%
	Subtotal 3 and 4	*10,482*	*10.7%*	*29.1%*
	Total	97,534	100.0%	100.0%

*Excludes 20% of vehicles that pass initial emissions test but fail initial visual or functional test.

The database we use for our analysis does not identify vehicles that exceed the cost repair limit without passing the test, and receive a waiver. Arizona DEQ reports that the waiver rate is about 4% of all vehicles that fail their initial test. If we assume that all of these waivered vehicles are classified as No Final Pass vehicles in our classification scheme, then the percentage of 1997 initial fail vehicles that never complete I/M testing is reduced to about 26%.

Another possibility for the high number of No Final Pass vehicles is that the VIN of a passing retest of these vehicles was entered incorrectly into the database, and therefore the passing retest was not matched with the initial test. To test this we sorted all tests of No Final Pass (including No Second Test) IM240 vehicles by vehicle license plate rather than VIN; it would be very unlikely for both the VIN and license plate to be incorrectly entered for the same vehicle. We found that only three of these vehicles had a subsequent retest with an invalid VIN; each of these vehicles failed the retest (one vehicle had two retests with invalid VINs, and failed both).

Tables 6 and 7 show the average initial and final emissions by I/M result, by vehicle type/class and for all vehicles. As noted above, we assume that the "final" emissions of vehicles with no second test, the Initial Pass and No Second Test vehicles, are the same as their initial emissions. IM240 emissions of the Final Pass vehicles are dramatically reduced by the I/M program: HC and CO emissions of these vehicles are reduced by over 60%, while NOx emissions are reduced by 45%. The percent reduction of CO and NOx emissions is somewhat greater for cars than light duty trucks. Presumably, much of this reduction is due to actual repairs made to vehicles; however, it is possible that initially failing vehicles can pass a retest without any repairs having been made. In addition, the emissions of No Final Pass vehicles also are reduced somewhat, presumably from partial repairs made to some vehicles in this group.

Table 6. Average IM240 Emissions and Percent Reduction by Vehicle Type and I/M Result, Unweighted by Annual VMT*

Type	I/M Result	Unweighted Average Emissions per Vehicle (adjusted grams per mile) HC Initial	HC Final	CO Initial	CO Final	NOx Initial	NOx Final	Percent Reduction HC	CO	NOx
Cars	1) Initial Pass	0.39	0.39	5.36	5.36	1.05	1.05	0.0%	0.0%	0.0%
	2) Final Pass	2.01	0.75	30.89	9.45	2.81	1.48	62.4%	69.4%	47.5%
	3) No Final Pass	2.95	2.71	43.59	39.69	2.63	2.48	8.1%	8.9%	5.9%
	4) No Second Test	3.04	3.04	46.46	46.46	2.55	2.55	0.0%	0.0%	0.0%
	Subtotal 3 and 4	*3.00*	*2.88*	*45.08*	*43.14*	*2.59*	*2.51*	*3.9%*	*4.3%*	*2.9%*
	All Cars	0.62	0.52	9.04	7.21	1.26	1.15	17.2%	20.3%	8.9%
LDT1	1) Initial Pass	0.60	0.60	8.83	8.83	1.41	1.41	0.0%	0.0%	0.0%
	2) Final Pass	3.21	1.22	43.12	16.61	3.68	2.09	62.2%	61.5%	43.1%
	3) No Final Pass	4.46	4.13	55.99	53.48	3.30	3.14	7.3%	4.5%	5.0%
	4) No Second Test	4.48	4.48	58.23	58.23	3.34	3.34	0.0%	0.0%	0.0%
	Subtotal 3 and 4	*4.47*	*4.28*	*56.98*	*55.58*	*3.32*	*3.23*	*4.0%*	*2.5%*	*2.8%*
	All LDT1	0.84	0.71	11.90	10.20	1.59	1.49	15.4%	14.3%	6.4%
LDT2	1) Initial Pass	0.72	0.72	10.26	10.26	2.02	2.02	0.0%	0.0%	0.0%
	2) Final Pass	4.26	1.45	55.77	21.77	3.81	2.37	66.0%	61.0%	37.8%
	3) No Final Pass	5.58	4.97	70.69	64.70	3.27	3.05	11.0%	8.5%	6.8%
	4) No Second Test	5.66	5.66	75.02	75.02	3.25	3.25	0.0%	0.0%	0.0%
	Subtotal 3 and 4	*5.62*	*5.31*	*72.83*	*69.81*	*3.26*	*3.15*	*5.5%*	*4.2%*	*3.5%*
	All LDT2	1.04	0.85	14.37	12.08	2.16	2.06	18.2%	16.0%	4.5%
All	1) Initial Pass	0.47	0.47	6.73	6.73	1.23	1.23	0.0%	0.0%	0.0%
	2) Final Pass	2.44	0.91	35.39	11.93	3.08	1.68	62.8%	66.3%	45.5%
	3) No Final Pass	3.29	3.02	46.58	42.82	2.76	2.60	8.1%	8.1%	5.8%
	4) No Second Test	3.38	3.38	49.58	49.58	2.70	2.70	0.0%	0.0%	0.0%
	Subtotal 3 and 4	*3.34*	*3.20*	*48.09*	*46.19*	*2.73*	*2.65*	*4.1%*	*4.0%*	*2.9%*
	Total	0.72	0.60	10.23	8.40	1.42	1.31	16.7%	17.9%	7.6%

*Excludes 4% of vehicles that pass initial emissions test but fail initial visual or functional test.

Table 7 shows that the CO emission reduction percentage of the loaded idle fleet (24%) is greater than the HC reduction (16%), and is greater than the CO reduction of the IM240 fleet (18%, Table 6). Emission reductions of Final Pass vehicles in the loaded idle fleet tend to be smaller than the percentage reductions of their counterparts in the IM240 fleet; however, because there are so many more Final Pass vehicles in the loaded idle fleet (Table 5 vs. Table 4), the result is larger overall emissions reductions across all vehicles.[7]

7. Idle emission reductions, both for the Final Pass vehicles and the overall fleet, are substantially higher than loaded idle emissions reductions. For instance, fleet idle emissions are reduced 26% for HC and 31% for CO.

Table 7. Average Loaded Idle Emissions and Percent Reduction by Vehicle Type and I/M Result, Unweighted by Annual VMT*

Type	I/M Result	Unweighted Average Emissions (ppm/%) per Vehicle				Percent Reduction	
		HC (ppm)		CO (%)			
		Initial	Final	Initial	Final	HC	CO
Class 3 (Cars and LDT1 with 4 or fewer cylinders)	1) Initial Pass	96	96	0.81	0.81	0.0%	0.0%
	2) Final Pass	190	104	2.41	0.90	45.1%	62.7%
	3) No Final Pass	243	232	2.76	2.67	4.4%	3.3%
	4) No Second Test	267	267	2.91	2.91	0.0%	0.0%
	Subtotal 3 and 4	*253*	*247*	*2.82*	*2.77*	*2.5%*	*1.8%*
	All Class 3	149	124	1.59	1.16	16.6%	26.7%
Class 4 (Cars and LDT1 with more than 4 cylinders	1) Initial Pass	85	85	0.77	0.77	0.0%	0.0%
	2) Final Pass	164	92	2.00	0.85	44.3%	57.5%
	3) No Final Pass	209	195	2.47	2.37	6.7%	4.1%
	4) No Second Test	210	210	2.26	2.26	0.0%	0.0%
	Subtotal 3 and 4	*209*	*202*	*2.38*	*2.33*	*3.7%*	*2.3%*
	All Class 4	117	98	1.24	0.95	16.4%	23.9%
Class 5 (LDT2)	1) Initial Pass	87	87	0.90	0.90	0.0%	0.0%
	2) Final Pass	155	94	2.03	1.04	39.2%	48.8%
	3) No Final Pass	210	194	2.35	2.24	7.4%	4.7%
	4) No Second Test	204	204	2.20	2.20	0.0%	0.0%
	Subtotal 3 and 4	*207*	*199*	*2.28*	*2.22*	*4.1%*	*2.6%*
	All Class 5	116	99	1.33	1.05	15.0%	20.9%
All	1) Initial Pass	87	87	0.81	0.81	0.0%	0.0%
	2) Final Pass	166	94	2.07	0.90	43.3%	56.5%
	3) No Final Pass	217	204	2.53	2.43	6.2%	3.9%
	4) No Second Test	222	222	2.40	2.40	0.0%	0.0%
	Subtotal 3 and 4	*219*	*212*	*2.47*	*2.42*	*3.4%*	*2.3%*
	Total	122	102	1.32	1.00	16.1%	23.7%

*Excludes 20% of vehicles that pass initial emissions test but fail initial visual or functional test.

Some of the differences in average emissions by I/M result is attributable to different vehicle age distributions in each of the vehicle groups. For instance, more newer vehicles are in the Initial Pass group, while more older vehicles are in the Final Pass or No Final Pass groups. Figures 3 through 5 present the average passenger car emissions by I/M result and model year for the IM240 fleet; Figures 6 and 7 present the same data for Class 4 vehicles of the loaded idle fleet. The initial emissions of the Initial Pass cars are compared with the initial and final emissions of the Final Pass and the No Final Pass (including No Second Test) groups.

The figures demonstrate that, for the most part, both initial and final HC and CO emissions are lower for newer vehicles than for older vehicles. This trend is due to a combination of better emissions control technology on newer vehicles, less aging and mileage accumulation of newer vehicles, and more stringent cutpoints for newer vehicles. (For example, the sharp decrease in HC emissions between model year 1990 and 1991 cars, most notable in for Final Pass and No Final Pass vehicles, is likely due to more stringent IM240 cutpoints applied to model year 1991 and newer vehicles.) Initial NOx emissions are fairly steady for 1990 and older cars; however,

Figure 3. Average HC by MY and I/M Result
Passenger Cars, 1997 Arizona IM240

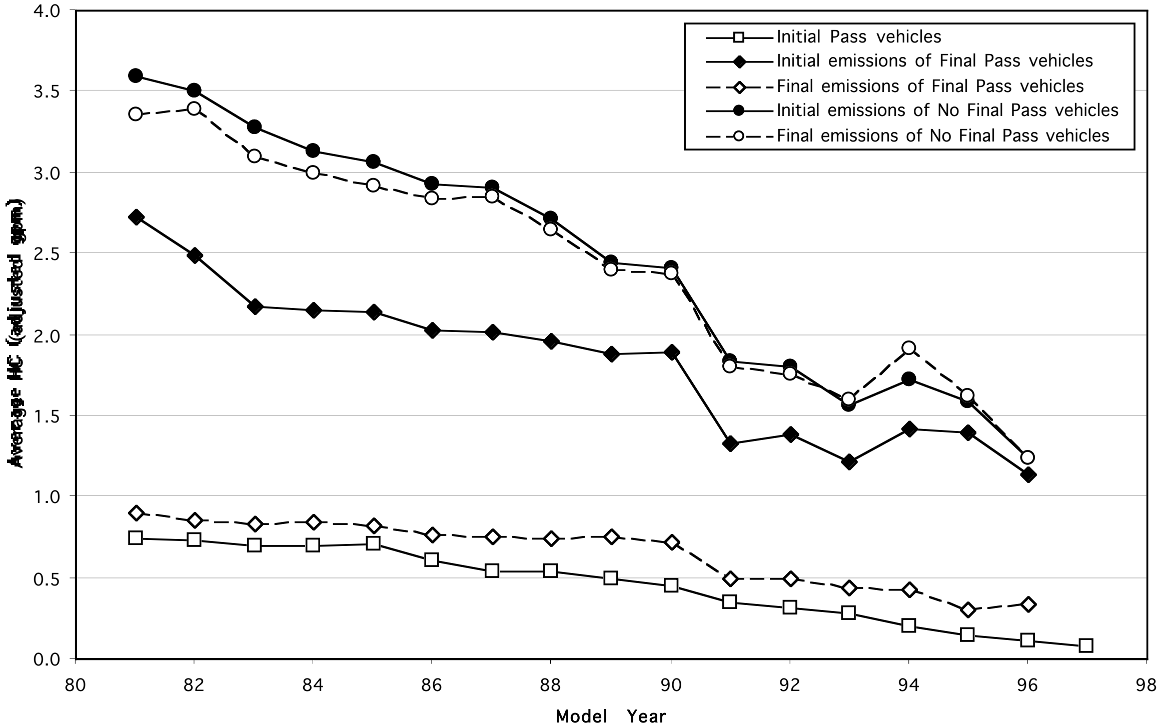


Figure 4. Average CO by MY and I/M Result
Passenger Cars, 1997 Arizona IM240

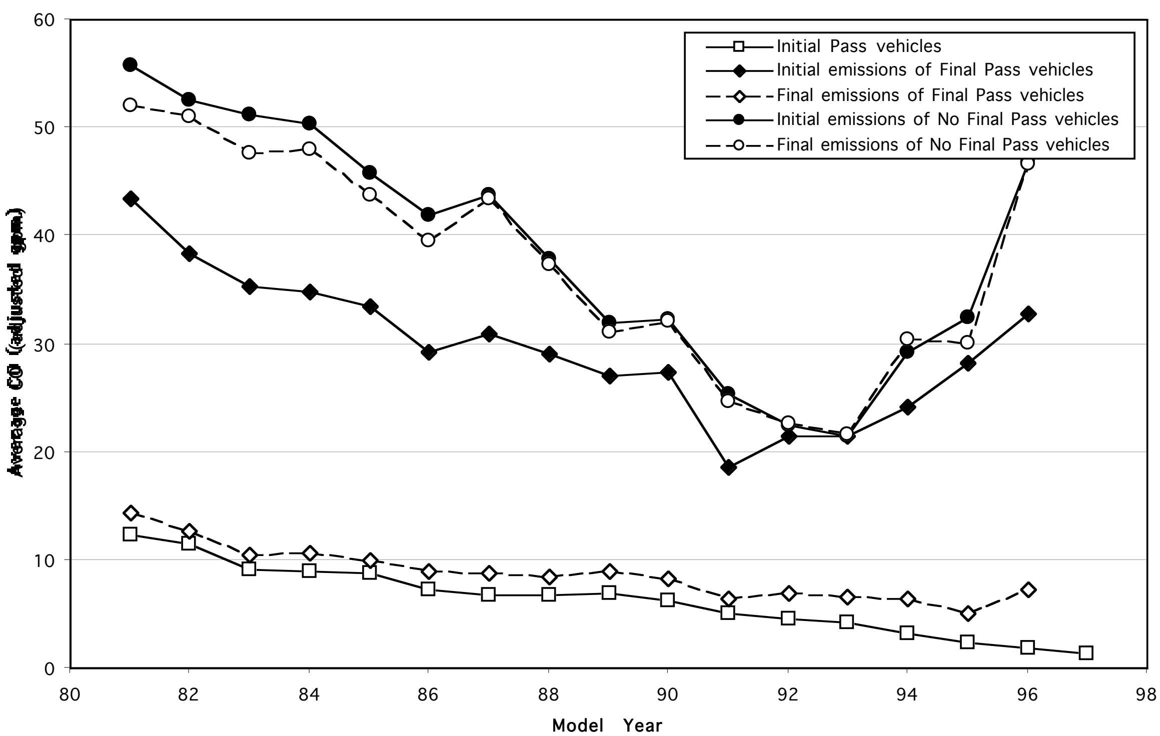


Figure 5. Average NOx by MY and I/M Result
Passenger Cars, 1997 Arizona IM240

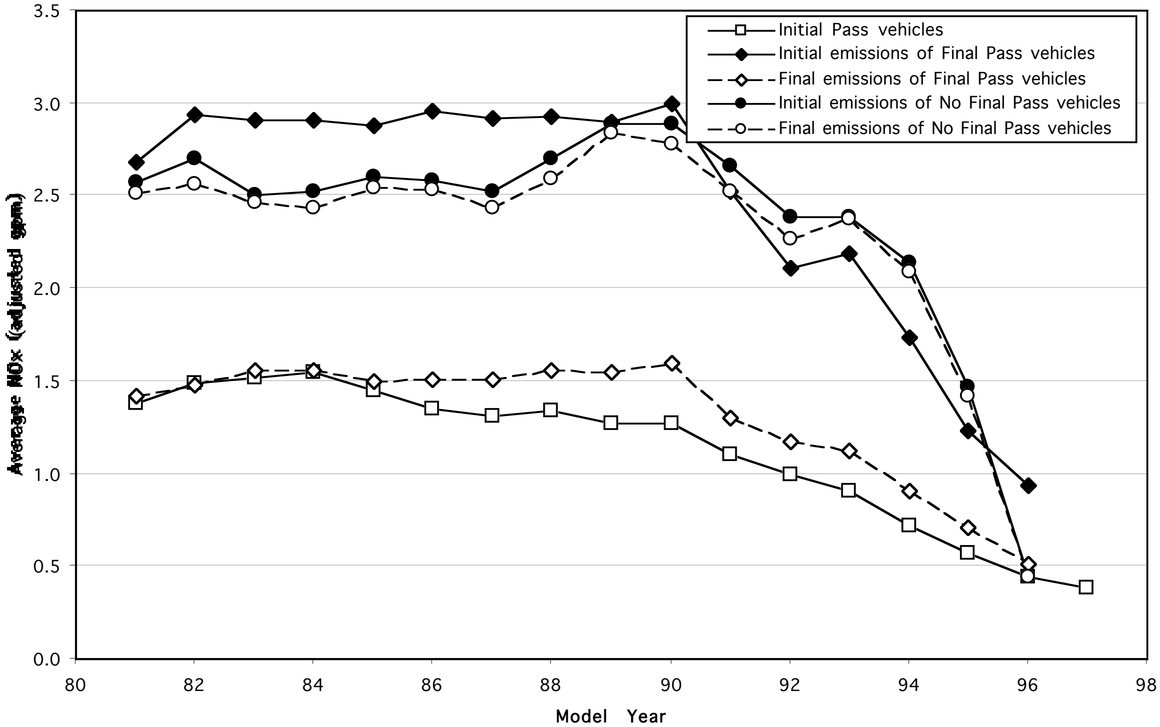


Figure 6. Average Loaded Idle HC by MY and I/M Result
Class 4 Vehicles, 1997 Arizona Idle

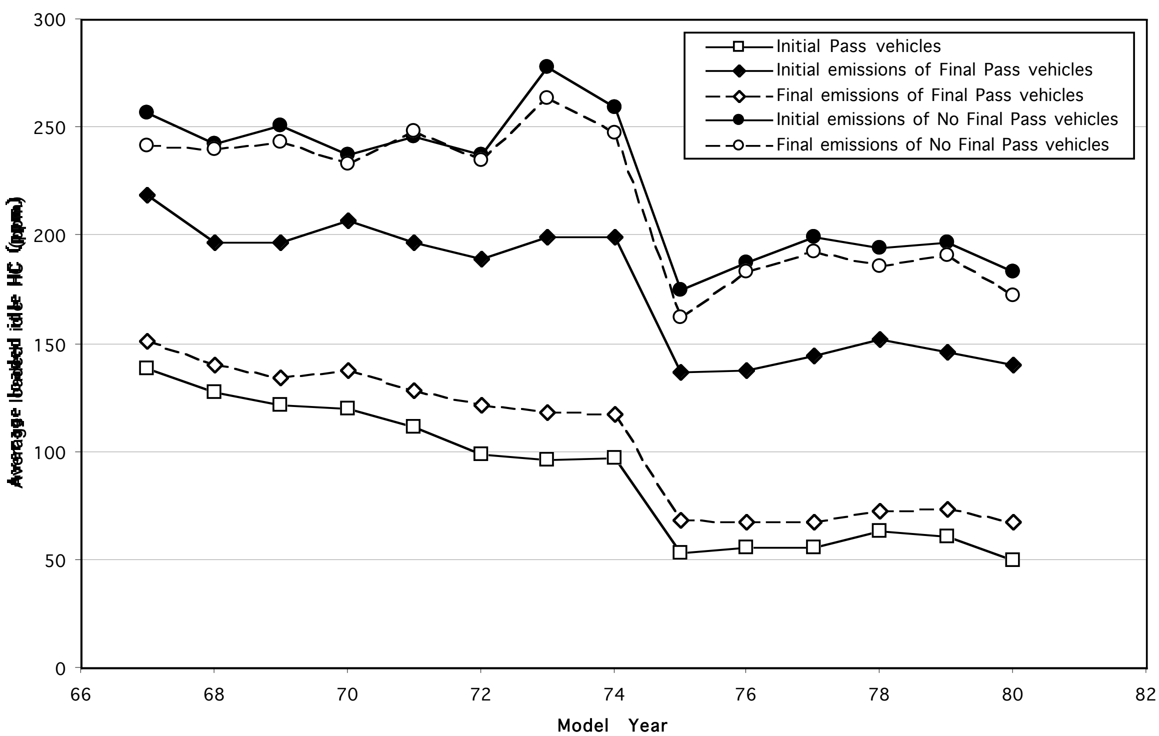
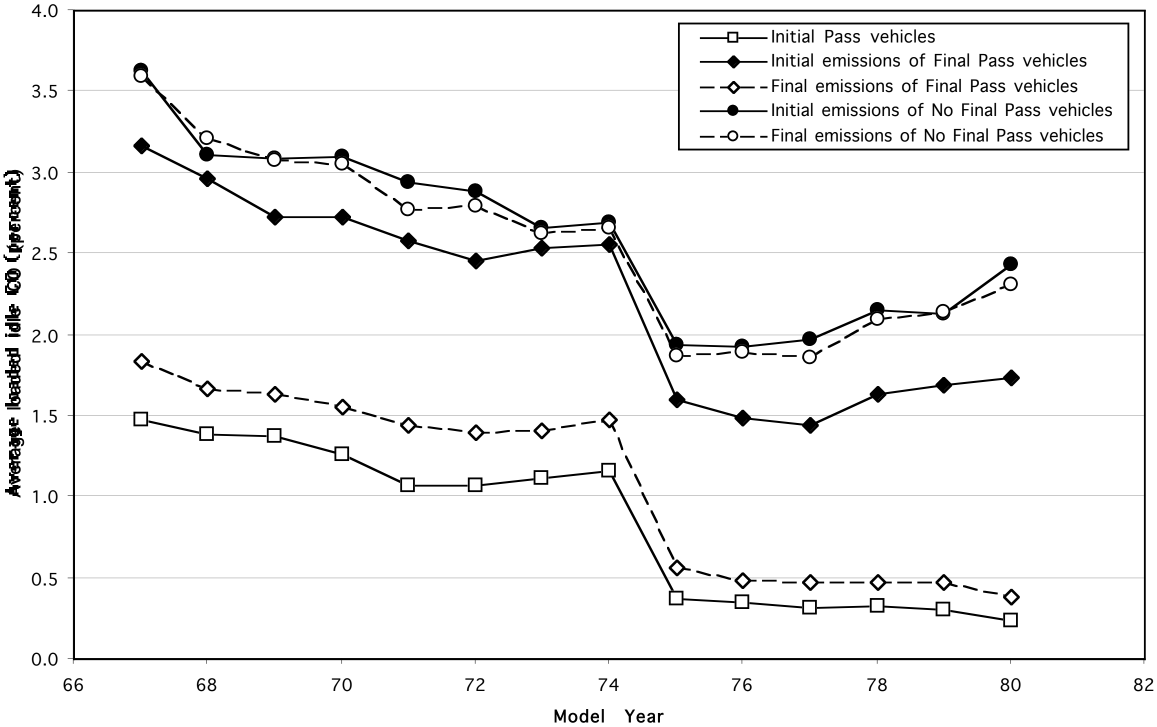


Figure 7. Average Loaded Idle CO by MY and I/M Result
Class 4 Vehicles, 1997 Arizona Idle

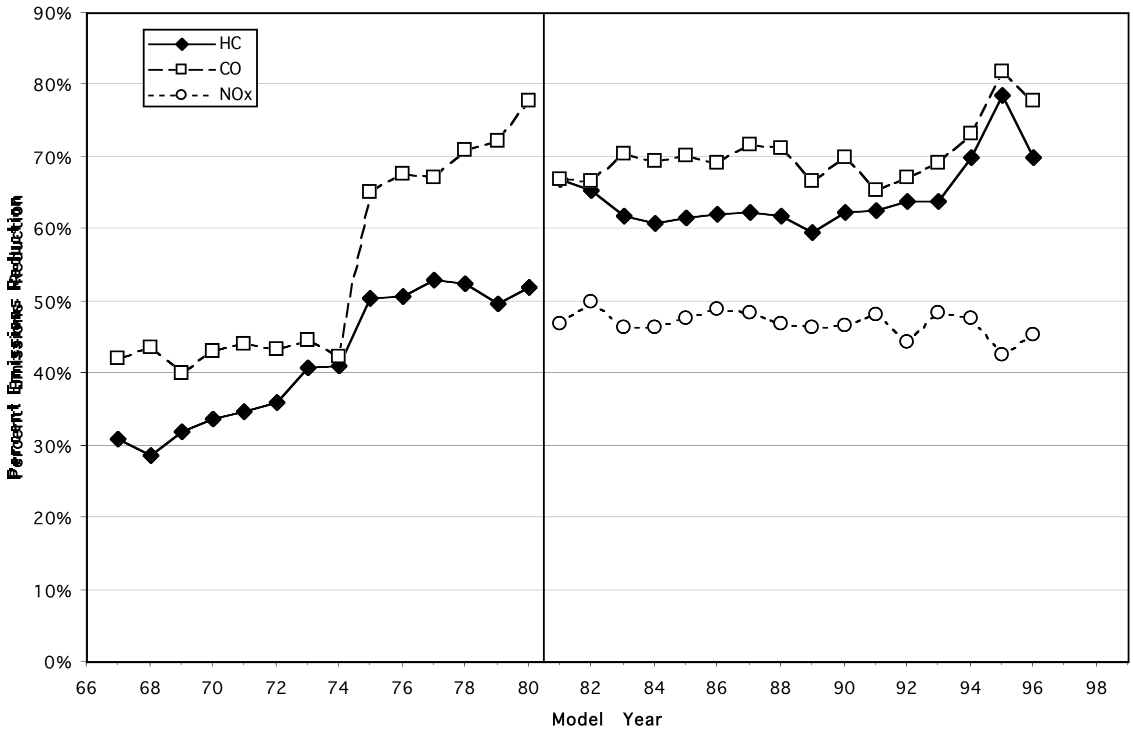


for 1991 and newer cars, NOx emissions are lower for newer cars. It is not clear why the trend in initial CO emissions of Final Pass and No Final Pass cars increases for 1993 and newer cars; this may be the result of out of state cars registering for the first time in Arizona (model year 1996 and newer vehicles already registered in the state were exempted from testing in 1997).

The figures show that Final Pass vehicle emissions are dramatically reduced by the program, at least as measured by program data. However, the emissions of Final Pass vehicles are not brought down to the level of emissions of Initial Pass vehicles. For the most part No Final Pass vehicles have higher initial and final emissions than Final Pass vehicles of the same age. However, older IM240 Final Pass vehicles have higher initial NOx emissions than older No Final Pass vehicles.

Figure 8 presents the percent emissions reduction for each pollutant, by model year, for Final Pass IM240 cars and loaded idle Class 4 vehicles. The figure indicates that the percentage emissions reductions of model year 1981 through 1993 Final Pass vehicles are fairly consistent by model year. HC and CO emission reduction percentages are slightly higher for 1993 and newer cars than for older cars. Percent reductions in loaded idle emissions are larger for model year 1975 through 1980 vehicles, than for older vehicles (loaded idle cutpoints are substantially stricter for 1975 and newer vehicles).

Figure 8. Percent Emissions Reduction by Model Year
Final Pass Cars/Class 4, 1997 Arizona IM240 and Loaded Idle



5. Estimating Effectiveness for the Entire I/M fleet

As discussed above, there are three major limitations of the I/M data that complicate any evaluation of the effectiveness of the overall Phoenix program:

1) One of two different emissions tests, the IM240 or the loaded idle test, is applied to each vehicle, depending on the vehicle's age. Each test measures vehicle emissions under different driving conditions, and reports emissions in different units. Therefore, emissions results as measured under the two tests are not directly comparable;

2) NOx emissions are not measured during the loaded idle test, therefore NOx emissions for the older fleet subject to loaded idle testing are not available; and

3) The loaded idle fleet is classified differently than the IM240 fleet, making it difficult to consistently weight emissions by annual vehicle miles traveled.

These limitations make it difficult to convert emissions concentrations from loaded idle testing into total mass emissions weighted by vehicle VMT, or the tons per day used for official emissions inventories and state implementation plans. In this section we attempt to determine the contribution of the loaded idle fleet to total I/M fleet emissions, and the tons of emissions reduced by the loaded idle program. We do this by extending the trend of the IM240 emission inventory by model year backward through model year 1967 vehicles, based on our analysis of the effectiveness of the program in reducing emissions of the loaded idle fleet.

Figures 9 and 10 show the trends in IM240 vehicles and their initial emissions in tons per day, respectively. Figure 9 demonstrates that the number of vehicles of all types increases as model year increases; the majority of the IM240 fleet is made up of relatively young vehicles. Figure 10 demonstrates a similar trend for NOx emissions; most of the NOx emissions come from the youngest vehicles. On the other hand, the peak of the HC and CO emissions distributions occurs around mid-1980s vehicles; fewer HC and CO emissions come from the youngest vehicles. These trends are due to the nature of HC and CO vs. NOx emissions. A few extremely high HC and CO emitters account for a relatively large portion of total HC and CO emissions, resulting in dramatically skewed distributions of HC and CO emissions. The range in NOx emissions is much smaller, resulting in a less skewed emissions distribution for NOx. Because NOx emissions are less skewed than HC or CO emissions, the number of vehicles heavily influences the NOx distribution in Figure 10.

Figure 9. Number of Vehicles by Type and Model Year
1997 AZ IM240

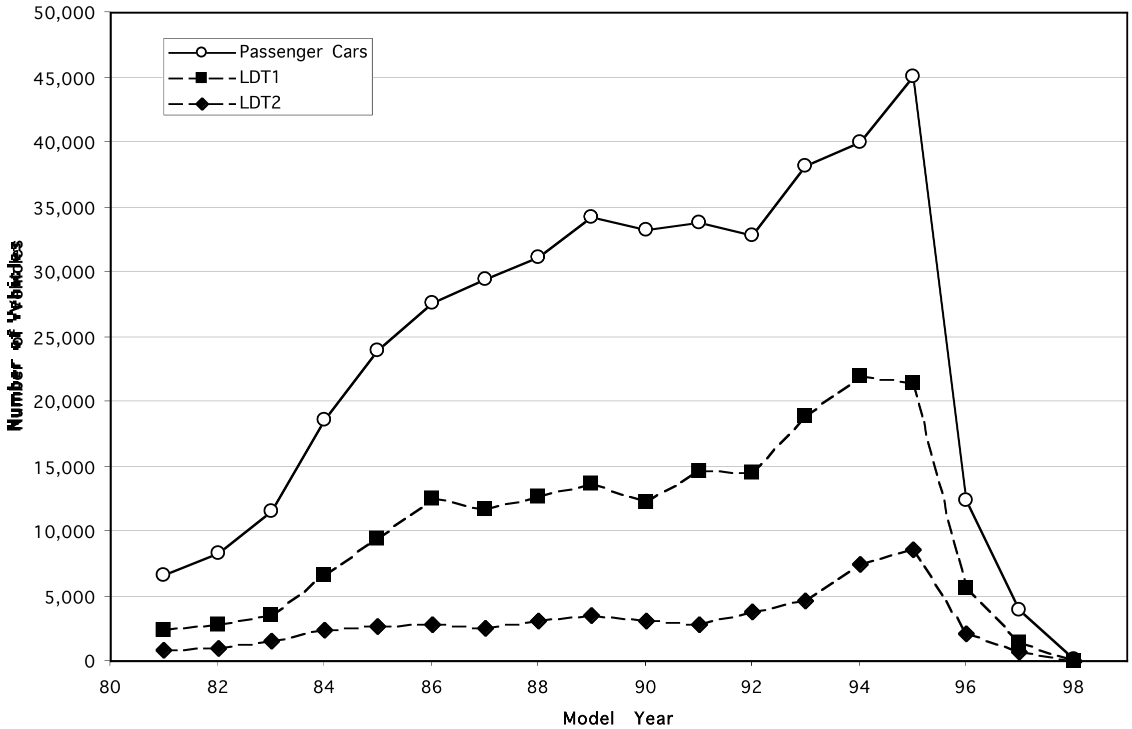


Figure 10. Total Emissions (tons per day) by Model Year

All Vehicle Types, 1997 AZ IM240

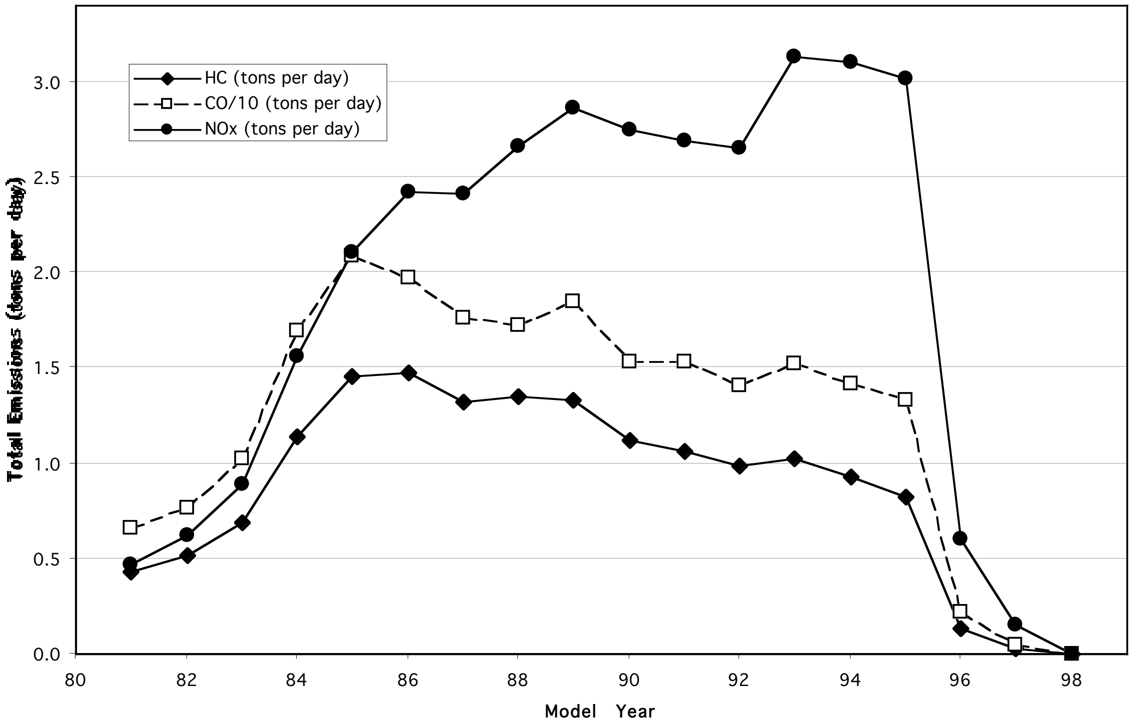
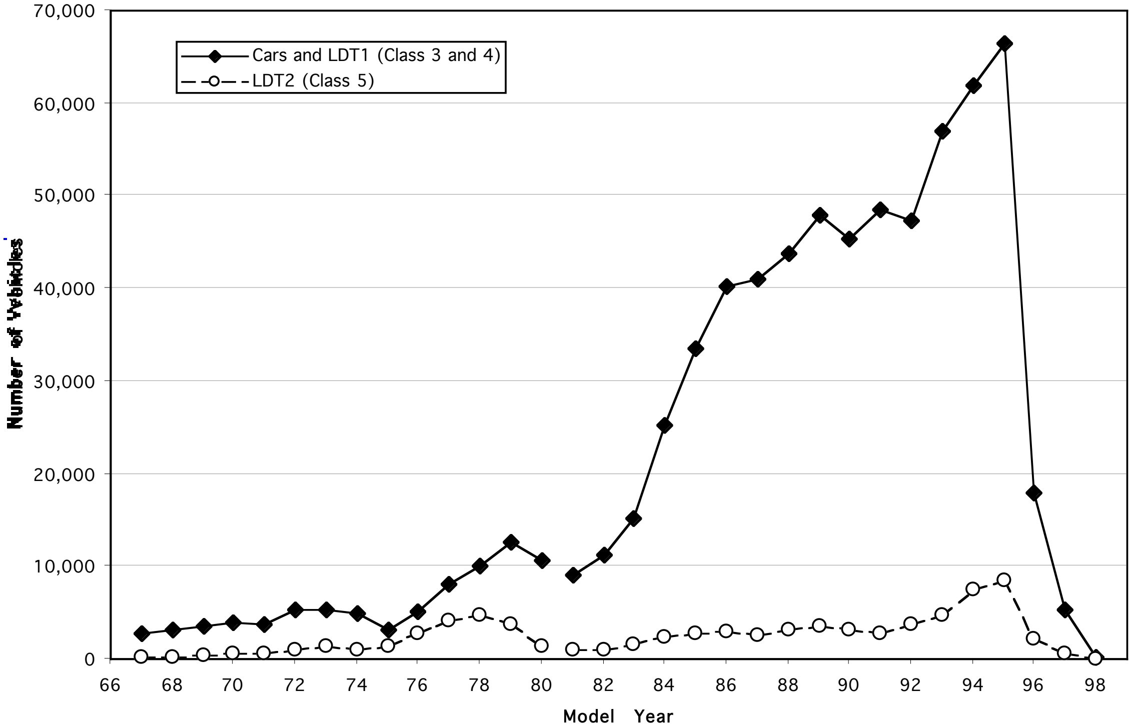


Figure 11 combines the distribution of IM240 vehicles in Figure 9 with the distribution of loaded idle test vehicles. Cars and LDT1 are combined into the same group to match the categories of the loaded idle test fleet (Classes 3 and 4). We see that there are many fewer loaded idle vehicles than IM240 vehicles. However, the vehicle distributions do not match perfectly; there are 20% more 1980 vehicles tested under the loaded idle program than 1981 vehicles tested under the IM240 program (17% more cars and LDT1, and 45% more LDT2). A possible explanation is that motorists perceive the IM240 test as more difficult to pass than the loaded idle test, and relocate their vehicles outside of the I/M area (either legally or illegally) to avoid the tougher IM240 test. However, this would not explain why the distribution of loaded idle vehicles peaks at model years 1978 and 1979, and declines for model year 1980 vehicles. The discrepancy between the number of LDT2 subject to the two tests is particularly disturbing; there are over five times as many model year 1978 LDT2 in the loaded idle fleet than 1981 LDT2 in the IM240 fleet. In fact, not until model year 1993 does the number of IM240 LDT2 approach the number of 1978 loaded idle LDT2.

Figure 11. Number of Vehicles by Type and Model Year
1997 Phoenix I/M Program



Our method to estimate the mass emissions of the loaded idle test vehicles involves taking the distribution of VMT-weighted emissions from these vehicles by model year and scaling it to the shape of the vehicle distribution. Calculating VMT-weighted loaded idle emissions is complicated since the loaded idle vehicles are not classified into cars and LDT1s. EPA's annual VMT assumptions by model year and type are dramatically different for older vehicles; for instance, estimated annual VMT for model year 1968 cars is nearly three times that of model year 1968 LDT1, while estimated annual VMT for model year 1980 cars is almost 40% higher than that of the same age LDT1. Using the car annual VMT weights for all Class 3 and 4 loaded idle vehicles results in an emissions inventory more than 40% greater than if the LDT1 weights are used for all Class 3 and 4 vehicles. We take the average of the car and LDT1 VMT weights for each model year to develop our VMT-weighted emissions for loaded idle vehicles.

Figures 12 through 14 show the new distributions of initial and final emissions by model year for both the loaded idle and IM240 vehicles. Since there are 20% more MY80 vehicles tested on the loaded idle than MY81 vehicles tested on the IM240, we scale the loaded idle emissions distribution so that the MY80 emissions in tons is 20% higher than the MY81 emissions. For NOx emissions from MY79 and older vehicles, we assume a smooth emissions distribution by model year where the emissions of each previous model year are 80% that of the next model year, with the constraint that MY67 vehicles account for 0.1 tons per day NOx. (The assumption of the smooth curve of NOx emissions underestimates the NOx contribution of model year 1979 and 1980 vehicles, but overestimates the contribution of 1975 and 1976 vehicles.) The initial and final emissions distributions by model year for HC, CO and NOx are shown in Figures 12, 13 and 14, respectively.

Figure 12. Estimated Total HC Emissions
(tons per day), by Model Year
1997 Phoenix I/M Program

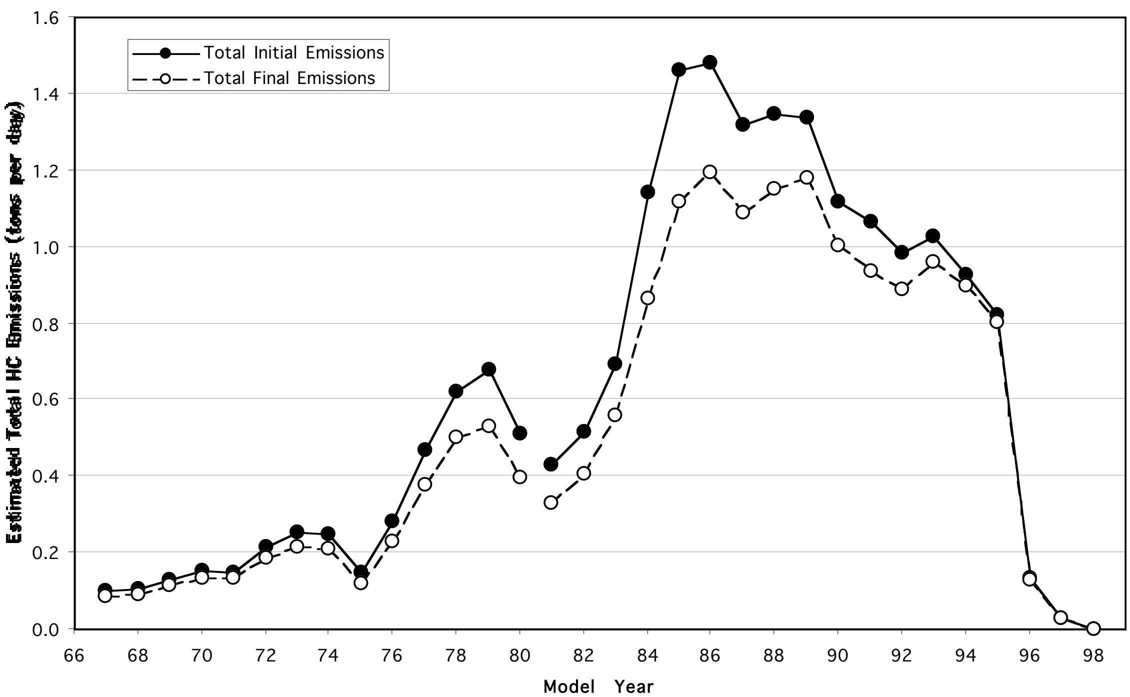


Figure 13. Estimated Total CO Emissions
(tons per day), by Model Year
1997 Phoenix I/M Program

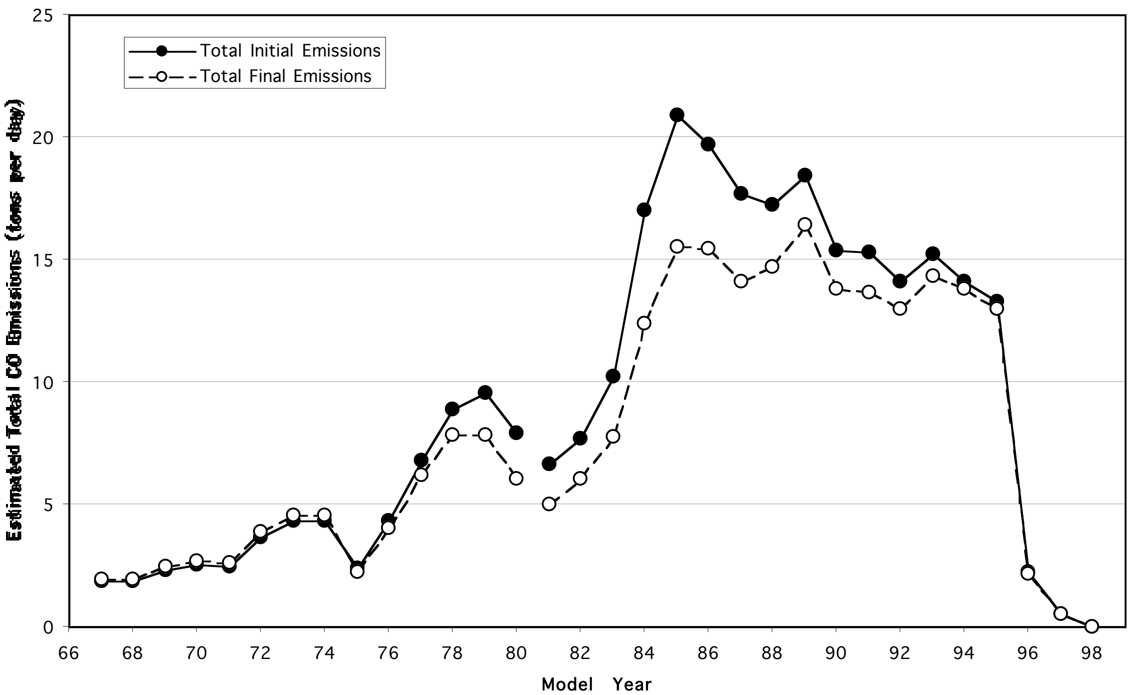
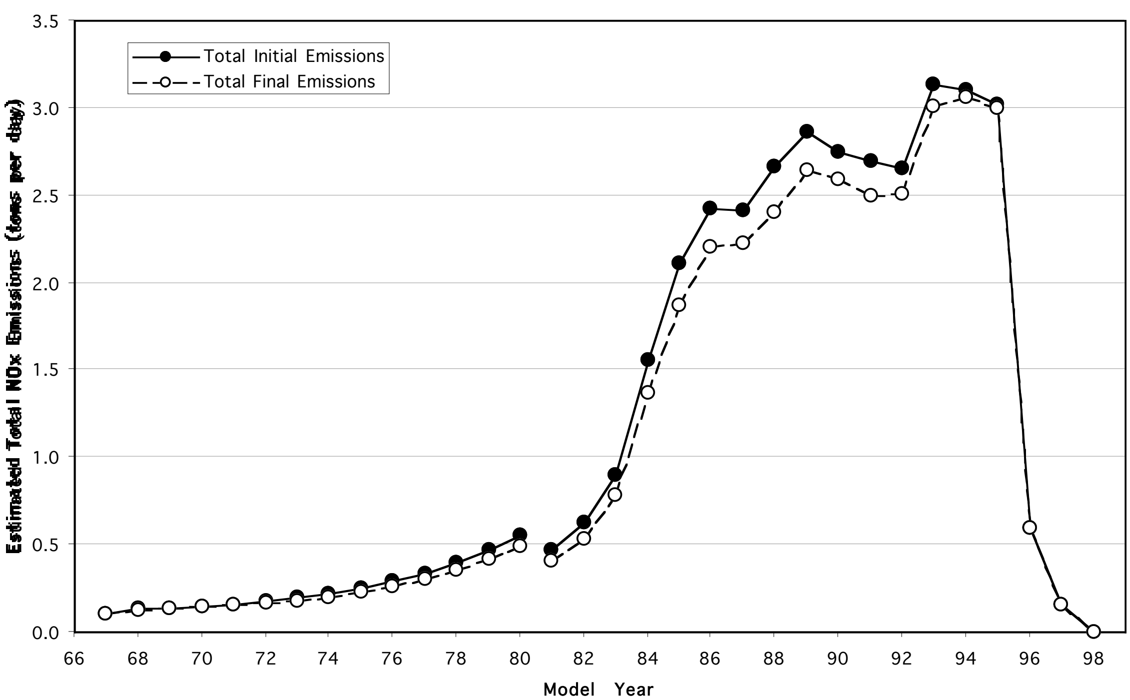


Figure 14. Estimated Total NOx Emissions
(tons per day), by Model Year
1997 Phoenix I/M program



The tons per day emissions and emission reductions derived from Figures 12 through 14 are shown in Tables 8 and 9. We estimate that the Phoenix I/M program reduces the emissions of the fleet reporting for I/M by 3.0 tons per day for HC, 38 tons per day for CO, and 2.6 tons per day for NOx. The majority of the estimated emissions reductions comes from the IM240 fleet:

76% for HC, and 88% for CO and NOx. The estimated percent reduction in VMT-weighted emissions is 15% for HC, 13% for CO, and 7% for NOx.

Table 8. Estimated Total Emissions by I/M Fleet, Tons per Day Weighted by Annual VMT

	Fleet	Number	HC (tpd)		CO (tpd)		NOx (tpd)	
			Initial	Final	Initial	Final	Initial	Final
Total Emissions	Idle	106,000	4.1	3.3	63.3	58.8	3.6	3.2
	IM240	670,768	15.8	13.5	225.6	191.7	34.1	31.9
	Total	776,768	19.9	16.9	288.9	250.4	37.7	35.1
Distribution of Emissions	Idle	14%	20%	20%	22%	23%	9%	9%
	IM240	86%	80%	80%	78%	77%	91%	91%
	Total	100%	100%	100%	100%	100%	100%	100%

Note: Absolute tons of emissions may not be comparable to official emissions inventories, due to conversion of fast pass/fast fail emissions to full IM240 emissions and exclusion of vehicles with invalid VINs, multiple initial tests, or that do not report for I/M testing.

Table 9. Estimated Emission Reductions by I/M Fleet, Tons per Day Weighted by Annual VMT

	Emission Reductions			Distribution of Emission Reductions			Percent Reduction		
Fleet	HC	CO	NOx	HC	CO	NOx	HC	CO	NOx
Idle	0.7	4.5	0.3	24%	12%	12%	18.0%	7.1%	8.8%
IM240	2.3	34.0	2.3	76%	88%	88%	14.4%	15.1%	6.6%
Total	3.0	38.5	2.6	100%	100%	100%	15.2%	13.3%	6.8%

Note: Absolute tons of emissions may not be comparable to official emissions inventories, due to conversion of fast pass/fast fail emissions to full IM240 emissions and exclusion of vehicles with invalid VINs, multiple initial tests, or that do not report for I/M testing.

6. Accounting for No Final Pass Vehicles

As noted above, about 26% of the vehicles that failed their initial IM240 test in 1997 never received a subsequent passing test through March 1998. It is possible that the program induced the owners of these vehicles to sell them or otherwise remove them from the I/M area. If so, the removal of these vehicles should be counted as a reduction in emissions attributable to the presence of the I/M program. However, if these vehicles are merely illegally re-registered outside of the I/M area, but continue to be driven regularly within the I/M area, the emissions of these vehicles must continue to be counted in the I/M area emission inventory. Whether or not these vehicles are still being driven in the I/M area does not affect estimates of emissions reductions on a per vehicle basis, as shown in Table 1. However, properly accounting for these vehicles will affect estimates of emissions reductions reported on an absolute tonnage basis, as presented in Table 3 (as well as Tables 8 and 9).

Table 3 assumes that all of the IM240 No Final Pass vehicles continue to be driven in the I/M area, and contribute to the “final” I/M emissions inventory. Table 10 assumes that none of these vehicles continue to be driven in the I/M area; these vehicles contribute to the “initial” I/M emissions inventory, but are removed from the “final” I/M emissions inventory. Removing all of the No Final Pass vehicles from the I/M area has a dramatic effect on the estimated effectiveness of the IM240 program, nearly doubling the percent reductions to 27% for HC and CO and to 11% for NOx, and the tonnage reductions to 4 tons per day for HC and NOx, and 61 tons per day for CO.

Table 10. Total Emissions and Percent Reduction, Weighted by Annual VMT (excludes all 1997 No Final Pass vehicles from final emissions)

Type	Number	Total Emissions (Tons per Day)						Percent Reduction		
		HC		CO		NOx				
		Initial	Final	Initial	Final	Initial	Final	HC	CO	NOx
Cars	414,173	8.7	6.0	125.7	82.5	18.8	16.1	31.2%	34.4%	14.2%
LDT1	182,608	5.1	4.0	72.9	59.8	10.6	9.6	20.7%	18.0%	9.1%
LDT2	52,904	2.0	1.6	26.9	22.0	4.8	4.5	21.0%	18.3%	5.5%
All	649,685	15.8	11.6	225.6	164.3	34.1	30.2	26.5%	27.2%	11.4%
Reduction	*21,084*		*4.2*		*61.3*		*3.9*			

Note: Absolute tons of emissions may not be comparable to official emissions inventories, due to conversion of fast pass/fast fail emissions to full IM240 emissions and exclusion of vehicles with invalid VINs, multiple initial tests, or that do not report for I/M testing.

Clearly a better understanding of the No Final Pass vehicles, and how many of them continue to be driven in the I/M area, is needed to properly estimate the effectiveness of the Arizona I/M program. In an earlier analysis we matched remote sensing data from 1996 and 1997 with 1995 and 1997 I/M test records (Wenzel, 1999b). About 30% of the 1995 No Final Pass (through March 1996) vehicles reported for their next scheduled I/M test in 1997. We compared the fraction of “1995 No Final Pass/tested in 1997” vehicles seen by remote sensing to the fraction of “1995 No Final Pass/not tested in 1997” vehicles seen by remote sensing. 7% of the fleet of vehicles reporting for testing in 1997 were seen by remote sensing over 2 years after their 1995 I/M test, while only 2% of the fleet that did not report for testing in 1997 were seen by remote sensing. The ratio of these two percentages (2% / 7%) gives us an estimate for the fraction of “1995 No Final Pass/not tested in 1997” vehicles still being driven in the I/M area: 27%.

If the fleet of vehicles initially tested in 1997 is similar to the fleet of vehicles initially tested in 1995, then we can assume that 30% of the 1997 No Final Pass vehicles will return for testing in 1999, and therefore will continue to be driven in the I/M area. In addition, of the 70% that will not report for testing in 1999, 30% will continue to be driven in the I/M area, or 20% (0.30 * 0.70 = 0.21) of all 1997 No Final Pass vehicles. Therefore, we estimate that about half (30% + 20%) of all 1997 No Final Pass vehicles continue to be driven in the I/M area more than 2 years after their 1997 I/M test. Table 11 shows the effect on total emissions and the percent reduction attributable to the I/M program, assuming that half of the 1997 No Final Pass vehicles continue to be driven in the I/M area. Under this assumption, the I/M program reduces HC and CO emissions by about 21%, and NOx emissions by about 9%; the tonnage reductions are 3 tons per day for HC and NOx, and 48 tons per day for CO.

Table 11. Total Emissions and Percent Reduction, Weighted by Annual VMT (excludes half of 1997 No Final Pass vehicles from final emissions)

Type	Number	Total Emissions (Tons per Day)						Percent Reduction		
		HC		CO		NOx				
		Initial	Final	Initial	Final	Initial	Final	HC	CO	NOx
Cars	422,635	8.7	6.7	125.7	92.8	18.8	16.7	23.3%	26.2%	10.9%
LDT1	184,247	5.1	4.2	72.9	62.3	10.6	9.8	16.8%	14.6%	7.6%
LDT2	53,346	2.0	1.7	26.9	22.9	4.8	4.5	17.6%	15.1%	4.6%
All	660,227	15.8	12.6	225.6	178.0	34.1	31.0	20.5%	21.1%	9.0%
Reduction	*10,542*		*3.2*		*47.6*		*3.1*			

Note: Absolute tons of emissions may not be comparable to official emissions inventories, due to conversion of fast pass/fast fail emissions to full IM240 emissions and exclusion of vehicles with invalid VINs, multiple initial tests, or that do not report for I/M testing.

Because of the limitations of the loaded idle test data, and the time constraints of this project, we have not estimated the effect of loaded idle No Final Pass vehicles permanently leaving the Phoenix area on the overall emissions inventory.

7. Other Issues

This analysis uses emissions test results from the Phoenix I/M program to evaluate the effect of the program in reducing vehicle emissions. The analysis compares the initial tests of vehicles with any subsequent tests to estimate emission reductions, both in terms of percent and in terms of tons of pollutants. A single year of I/M program data can give an indication of the initial effectiveness of vehicle repairs performed under the program. However, there are several limitations with basing a program evaluation solely on emissions test results from the program itself:

- The emissions difference between the initial and final tests does not capture all of the emissions reductions that the program may be causing; dirty vehicles may leave the area, motorists may take better care of their vehicles, and motorists may pay more attention to purchasing cleaner vehicles as a result of the I/M program.

- Some of the emissions difference between the initial and final tests may not be due to repair at all. For example, more extensive preconditioning can cause a failed vehicle to pass a retest without repairs being made. Or a vehicle may pass a retest when environmental conditions (ambient temperature and humidity) are more favorable. Or the effect of regression to the mean may cause a moderately high emitter to have slightly lower emissions on a retest and pass. These are three of several possible explanations for why the difference between the initial and final readings may be overestimating the amount of emissions reduction.

- In-program data measure the effectiveness of any vehicle repairs immediately after such repairs have been made. In effect, such an analysis assumes that all repairs made remain effective. However, repaired components on some vehicles may fail shortly after testing, or the repair may not address the underlying cause of the higher emissions. Evaluations based on in-program data do not account for the effect of insufficient, or temporary, repair of vehicles.

- As discussed above, the presence of an I/M program may induce some owners to register their vehicles outside of the I/M area, particularly if they suspect their vehicle will fail an I/M test. If these vehicles are indeed high emitters, and they are legitimately registered outside of the I/M area (and no longer driven in the I/M area), then area emissions will have been reduced. However, if these high emitter vehicles were re-registered merely to avoid I/M testing, and continue to be driven in the I/M area, area emissions will be unchanged. Evaluations using in-program data cannot account for whether vehicles re-registered outside of the program area are high emitters, and what fraction of them continue to be driven in the I/M area, contributing to area emissions inventories.

- Because the I/M test is scheduled, drivers may make temporary repairs or adjustments to vehicles immediately prior to testing. If these repairs result in permanent emissions reductions, in-program data will underestimate the effect of the program in reducing these emissions. If these are merely adjustments made to pass the I/M test, with the vehicles readjusted after passing, program data will correctly measure the percent emissions reduction (none) but will underestimate total fleet emissions.

Some of these issues can be addressed using multiple years of program data. For instance:

- The long-term effectiveness of repairs made to vehicles can be determined by tracking individual vehicles participating in the program over several test cycles. An earlier analysis of 1995 and 1997 data from the Arizona I/M program indicates that 37% of the vehicles that failed their initial test in 1995 but passed a subsequent retest failed their next regularly-scheduled test in 1997. The repeat failure rate ranges from under 15% for newer vehicles to nearly 45% for the oldest vehicles. Of the vehicles that failed in both years, about half failed for the same combination of pollutants in both years, suggesting that, for many vehicles, the repairs made in 1995 did not address the underlying causes of high emissions (Wenzel, 1999a).

- Individual vehicles that are not tested in subsequent I/M test cycles (either due to registering outside of the I/M area, or to otherwise avoiding the I/M program) can be identified. Vehicles that have migrated into the I/M program can also be identified, and their emissions compared with those that have participated in the program. The earlier analysis found that 40% of all vehicles tested in 1995 did not return for testing in 1997. The vehicles that did not report for testing in 1997 tended to be older, and have higher emissions, than the vehicles that did reported for testing in both years. Similarly, about half of the vehicles that were tested in 1997 were not tested in 1995. Of these not tested in 1995, half were either: MY94 and older out of state cars newly registered in Arizona (23%); MY95 cars exempted from testing in 1995 (18%); or MY96 and newer cars voluntarily tested in 1997 (8%). The vehicles tested in 1997 but not in 1995 tended to have higher emissions than the vehicles tested in both years.

On-road emissions testing, either using remote sensing data or roadside testing of vehicles randomly pulled over, can also be used to address some of these issues. In particular, on-road emissions testing can be used in two ways to evaluate I/M program effectiveness:

1) On-road emissions testing programs measure vehicles at different times relative to their last I/M test. Therefore these data can be used to estimate how quickly repair effectiveness diminishes over time, as well as how much repair is made just prior to the I/M test (Wenzel, 1999b).

2) Remote sensing programs measure almost every vehicle that drives by the instrument, regardless of whether it is participating in the I/M program. Remote sensing data therefore can be used to estimate the number and emissions of vehicles legally exempted from, or illegally avoiding, the I/M program, as well as estimating their emissions. In addition, remote sensing data can identify individual vehicles that never complete the current I/M cycle, or that do not report for testing in a subsequent test cycle, but are still being driven in the I/M area.

8. Summary

In this report we use emissions test result data from 1997 to evaluate the effectiveness of the enhanced I/M program in reducing vehicle tailpipe emissions in Phoenix, Arizona. Because the program requires a loaded idle, rather than IM240, test for 1980 and older vehicles, we analyze the effectiveness of the program on the two fleets of vehicles separately. The analysis does not consider the effect of the I/M program on reducing evaporative HC emissions. Because Arizona allows vehicles to fast pass or fast fail the IM240 test, we must convert IM240 "short test" results to full IM240 test equivalents. The relatively simple method we use to make this conversion is comparable to other more detailed methods.

Comparison of initial and final IM240 tests indicates that the program is reducing the average per vehicle emissions by 16% for HC, 17% for CO, and 7% for NOx, for the entire vehicle fleet. After weighting per vehicle emissions by estimated annual miles traveled, the fleetwide emissions reductions are 2.3 tons per day (14% reduction) for HC, 34 tons per day (15% reduction for CO), and 2.3 tons per day (7% reduction) for NOx. CO and NOx reductions appear to be substantially larger for cars than for light duty trucks. Per vehicle emissions of the loaded idle fleet are reduced by 15% for HC and 23% for CO.

About 11% of all vehicles fail their initial IM240 emissions test; the failure rate is slightly higher for passenger cars (12%) than for light duty trucks (8%). The initial failure rate for the loaded idle test is 37%. Of the vehicles that fail their initial test, only 70% received a final passing test through March 1998; 30% did not receive a final passing test through March 1998. Because waivered vehicles are not identified in the data, the actual percentage of No Final Pass vehicles is likely to be closer to 26%. The percentage of No Final Pass cars is greater than the percentage of No Final Pass trucks.

The percent reductions in loaded idle emissions for Final Pass vehicles tend to increase by model year, with larger reductions for newer vehicles. There is a large increase in percent reduction for model year 1974 through 1980 vehicles, presumably due to stricter cutpoints applied to those vehicles. The percentage reductions of IM240 Final Pass vehicles from model years 1981 through 1993 are fairly constant by model year. HC and CO emission reduction percentages tend to increase after model year 1993.

We use a relatively crude method to estimate total emissions and emission reductions in tons per day for the loaded idle fleet, in order to estimate the tonnage reductions for the entire Phoenix I/M program. We estimate that the program reduces the emissions of the fleet reporting for I/M by 3.0 tons per day for HC, 38 tons per day for CO, and 2.6 tons per day for NOx. The majority of the estimated emissions reductions comes from the IM240 fleet: 76% for HC, and 88% for CO and NOx. The estimated percent reduction in total emssions is 15% for HC, 13% for CO, and 7% for NOx.

The estimated effectiveness of the I/M program depends on whether the No Final Pass vehicles have been permanently removed from the I/M area, or if they continue to be driven in the I/M area. The effectiveness of the program on the IM240 fleet nearly doubles if one assumes that all IM240 No Final Pass vehicles have been permanently removed from the area. Analysis of 1995 IM240 test data and remote sensing data indicate that about half of the No Final Pass vehicles continue to be driven in the I/M area. If this information is correct for vehicles tested in 1997, the 1997 I/M program resulted in a 22% reduction in HC and CO, and a 9% reduction in NOx from the IM240 fleet. These percentage reductions are equivalent to 3.0 tons per day for HC and NOx, and 48 tons per day for CO.

Analysis of a single year of I/M program test data can only provide a partial understanding of the program's effectiveness in reducing emissions. Tracking of individual vehicles over several I/M cycles can reveal important information on long-term effectiveness of vehicle repair, and changes in the fleet reporting for I/M testing. In addition, an independent source of on-road emissions tests, such as from a remote sensing measurement program, can provide additional information on repair effectiveness, the effect of pre-test repairs on emissions, and the number and emissions of vehicles avoiding the I/M program.

9. References

Acurex Environmental Corporation. 1997. *Update of Fleet Characterization Data for Use in MOBILE6*. May 16.

T. Wenzel. 1999a. "Evaluation of Arizona's Enhanced I/M Program," presentation at the 9th CRC On-Road Vehicle Emissions Workshop, San Diego, CA, April.

T. Wenzel. 1999b. "Human Behavior in I/M Programs," presenation at the 15th Annual Mobile Sources/Clean Air Conference, Snowmass, CO, September.

APPENDIX R

DRAFT LBNL-41452

ANALYSIS OF EMISSIONS DETERIORATION OF IN-USE VEHICLES, USING ARIZONA IM240 DATA

Tom Wenzel, Lawrence Berkeley National Laboratory
Marc Ross, University of Michigan
Robert Sawyer, University of California at Berkeley

Presented at the
Society of Automotive Engineering Government/Industry Meeting
Washington, DC
May 5-7, 1997

Energy Analysis Program
Environmental Energy Technologies Division
Lawrence Berkeley National Laboratory
University of California
Berkeley, CA 94720

This work was supported by the Office of Mobile Sources, Office of Air and Radiation, of the U.S. Environmental Protection Agency. Prepared for the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

Abstract

Computer models that forecast in-use vehicle air pollutant emissions (i.e. MOBILE and EMFAC) base assumptions of how vehicle emissions increase as they age on limited laboratory emissions tests. States' adoption of dynamometer-based testing as part of enhanced vehicle emission inspection and maintenance (I/M) programs has resulted in in-use emissions data on hundreds of thousands of vehicles. This paper analyzes how emissions increase as vehicles age, using data from the Arizona I/M program

Arizona I/M Data

Arizona is one of the first states to use the Enhanced I/M testing procedure recommended by EPA. The procedure involves testing vehicles on a treadmill-like device, called a dynamometer, which simulates the vehicle driving a 240-second speed-time trace, called the IM240, that tests emissions under varying vehicle operating conditions. The procedure was developed to simulate the detailed emissions testing manufacturers perform on new vehicles.

This analysis is based on initial IM240 tests conducted on in-use passenger cars in Arizona in 1995. Arizona's program allows the cleanest vehicles to pass inspection after only 30 seconds of testing (fast passes), and the dirtiest vehicles to fail after 94 seconds of testing (fast fails). Other vehicles can pass or fail at any time before the full 240 second test is completed. A random 2% sample of vehicles are given the full 240-second test, regardless of whether they pass or fail the test. In addition, a small number of vehicles are tested over the entire test without fast-passing or fast-failing. We included the fast pass/fast fail tests in our analysis.

There are two major drawbacks with using the fast pass/fast fail data. First, vehicles are tested over different portions of the IM240 cycle, resulting in inconsistent emissions values. This is particularly important for vehicles passed immediately after 30 seconds of testing, for two reasons: these vehicles represent from 40 to 60 percent of all cars of a given model year, and a given vehicle will have substantially higher gram per mile emissions at second 30 than at second 240. The second drawback is that long vehicle wait times may affect emissions measurements. The engines and catalysts of vehicles that wait 15 or more minutes prior to testing may have cooled down sufficiently, resulting in higher emitssion than if they were properly warmed up (or "preconditioned") prior to testing (Heirigs and Gordon, 1996). Consequently, inconsistent preconditioning of vehicles may overstate average gram per mile emissions.

However, there are benefits to using the fast pass/fast fail data, rather than limiting the analysis to the random sample of full IM240 tests. The sheer number of tests allow detailed analyses of emissions (for instance, by model year and mileage), without losing statistical significance. And the full dataset is likely more representative of the Arizona on-road fleet than the random sample.

Methodology

We made several refinements to the data to improve our analysis. First, we adjusted the reported test results to more accurately reflect results if each vehicle was tested on a full IM240. The contractor reports test results as total grams divided by the distance of the full IM240 test (1.96

miles). Our adjustment involved two steps: calculating actual grams per mile based on actual miles each vehicle was driven, and correcting for different test durations. We obtained from Ontario average second by second emissions data from 11,000 vehicles tested over the full IM240. Figure 1 shows the speed time trace of the IM240 (right scale), and the average gram per mile emissions fo the Ontario test fleet at each second of the test (left scale). For each second of the test, cumulative grams are divided by cumulative miles for each vehicle, and the results are averaged over the fleet. The highest average gram per mile values occur at second 30, and decrease as the test continues. The hardest acceleration in the IM240 occurs just before second 160; this

Figure 1. Average gpm Emissions at Each Second of IM240, Ontario Data (n=11,000 full IM240 tests)

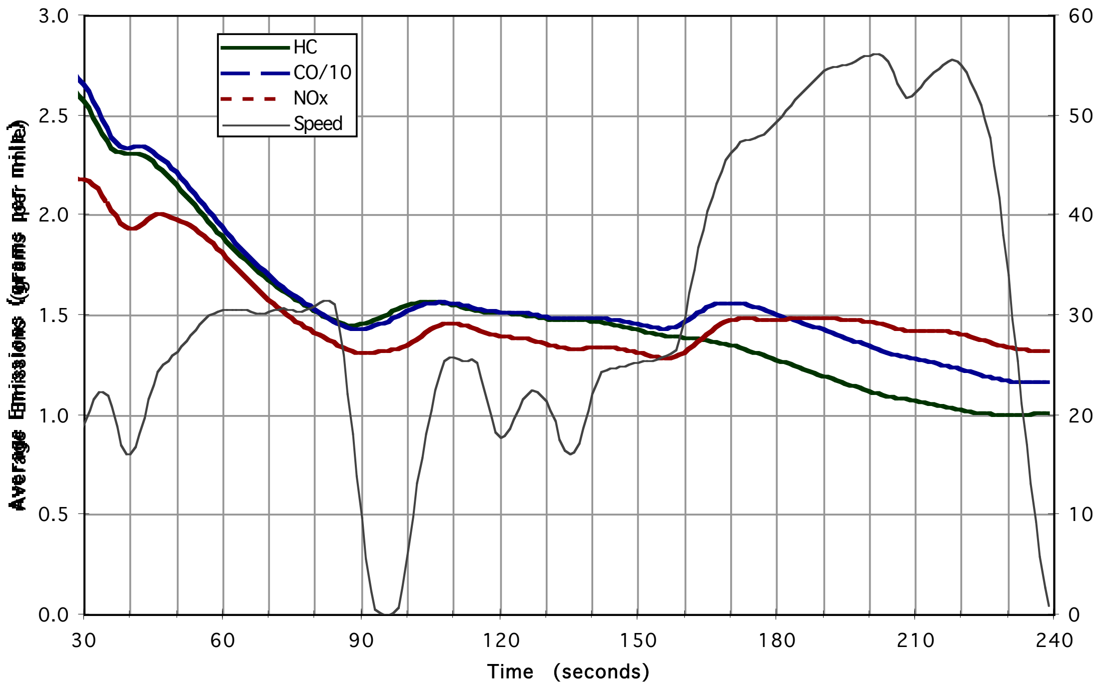
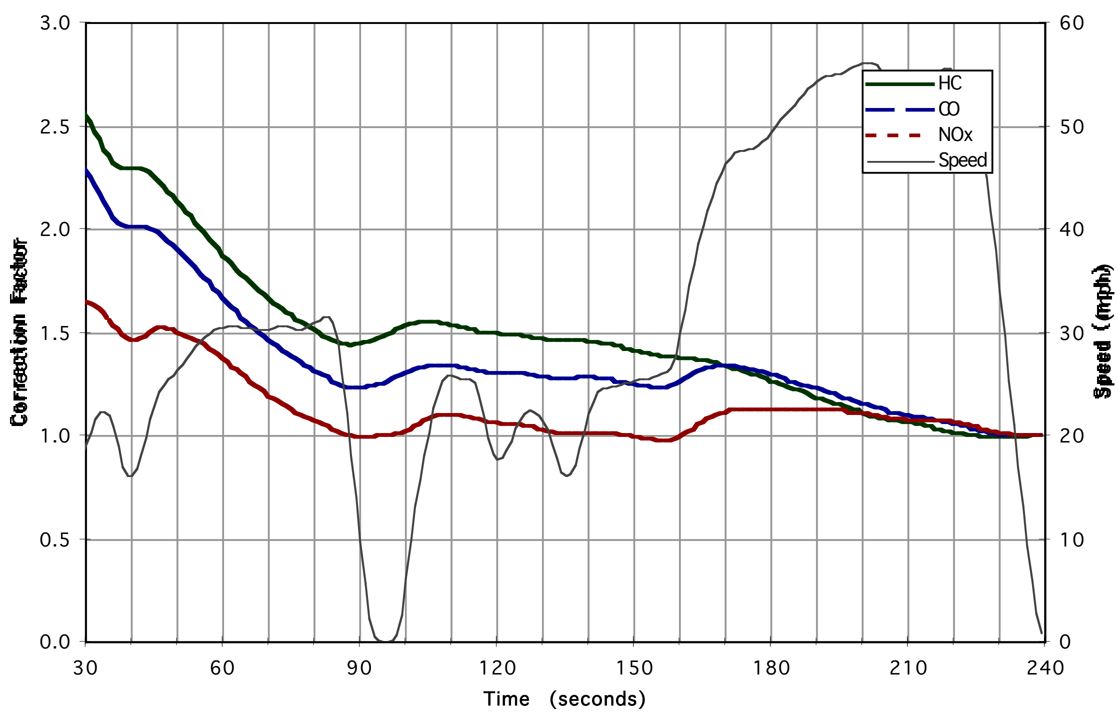


Figure 2. Emission Correction Factor for Each Second of IM240, Ontario Data (n=11,000 full IM240 tests)



acceleration causes the cumulative average gram per mile values for CO and NOx to increase slightly.

We then developed a correction factor for each second of the test, for each pollutant, based on the ratio of the average emissions at each second to the average emissions for the full IM240. As shown in Figure 2, the correction factors are quite large for cars passed immediately after 30 seconds; for these cars we divided measured gram per mile values by 2.5 for HC to obtain full-IM240 equivalent emissions.

Figures 3 through 5 show the effect of our two adjustments on the test results reported by the contractor. The dashed lines represent converting the reported results into gram per mile values, and the heavy lines represent the downward adjustment of the gram per mile values to account for different test durations. In general, our adjustments substantially increase the reported emissions from cars tested on shorter portions of the IM240 (for example, emissions from cars passed after 30 seconds of testing are increased by a factor of 4 to 7, depending on the pollutant). Our adjustments resulted in smaller increases from reported emissions from cars tested over longer segments of the IM240.

We made no correction for the problem of inconsistent preconditioning, although we did attempt to identfy individual cars that may not have been properly warmed-up prior to testing. Arizona allows cars that fail the test a second chance to pass, based on the emissions over the second half (Phase 2) of the IM240, when the car has presumably been sufficiently warmed-up. We assumed that cars that failed the composite cutpoints, but passed the Phase 2 cutpoints, were not fully warmed-up; these cars represent less than 5 percent of the fleet tested.

Figure 3. Average HC Emissions By Test Duration, MY83-94 Cars
1995 AZ IM240 (n=366,000)

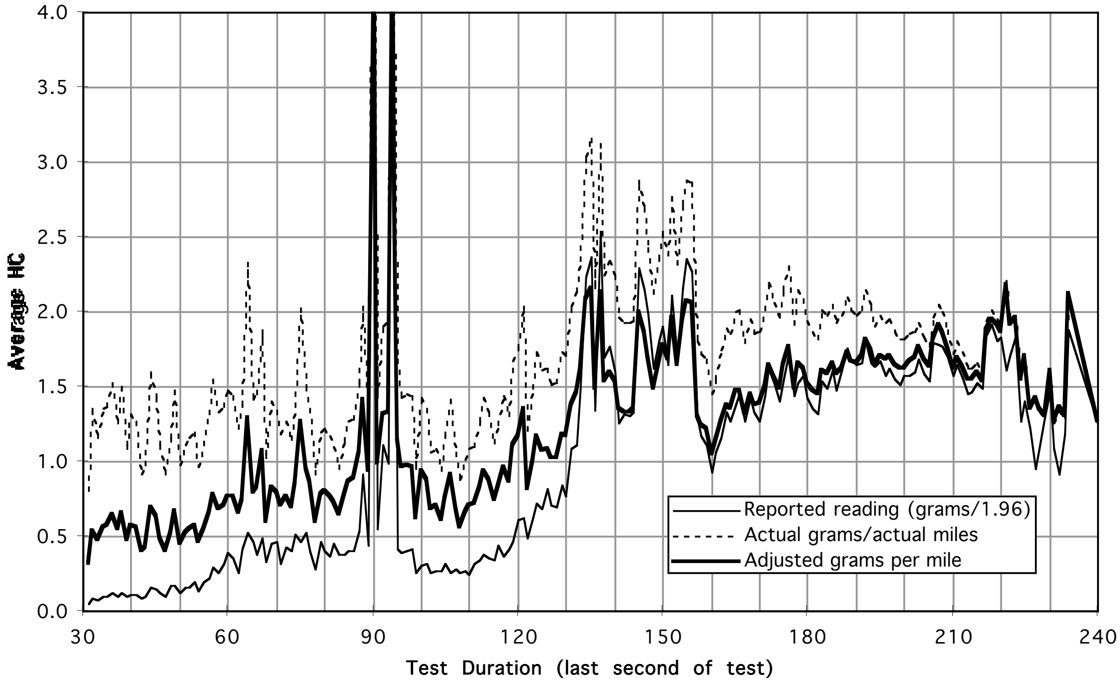


Figure 4. Average CO Emissions by Test Duration, MY83-94 Cars
1995 AZ IM240 (n=366,000)

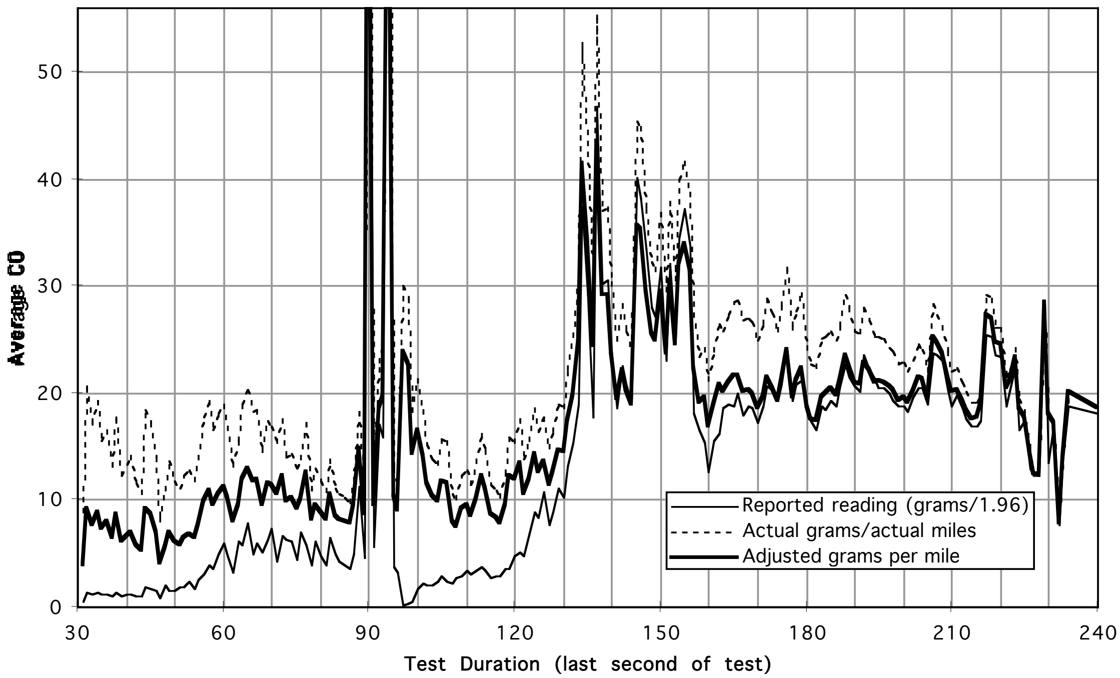


Figure 5. Average NOx Emissions by Test Duration, MY83-94 Cars
1995 AZ IM240 (n=366,000)

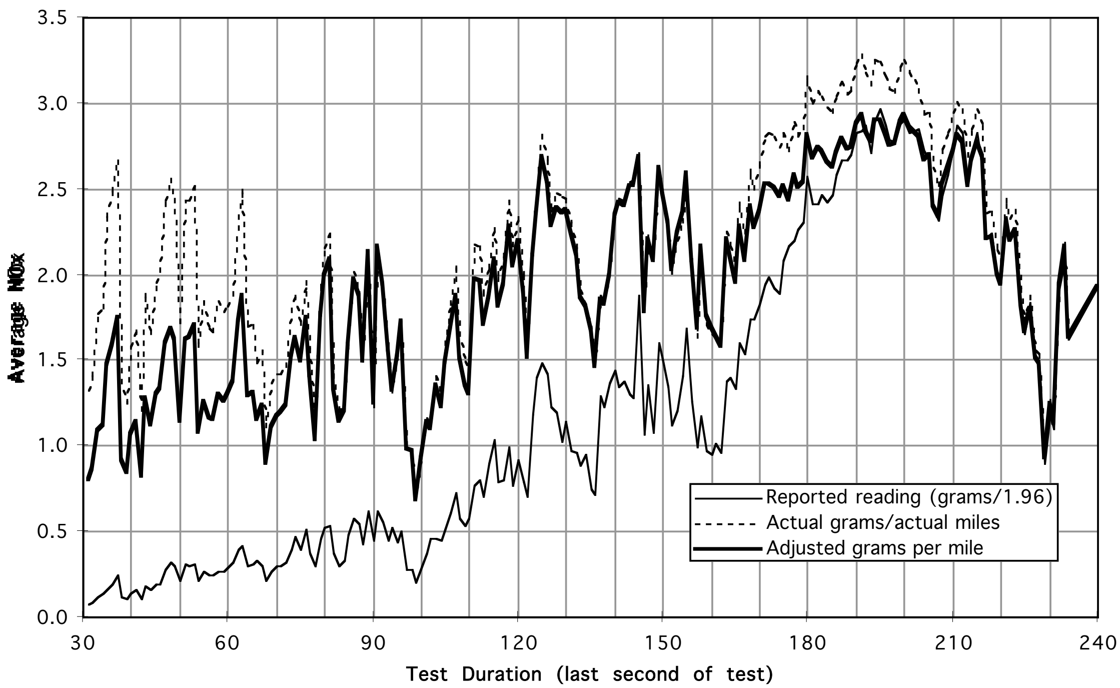


Figure 6 shows the distribution of cars by odometer reading. We found a large number of vehicles with zero and 100,000 mile odometer readings. We assumed these were misread odometers, and removed these cars from the analysis. That left us with the initial IM240 tests of 350,000 cars tested in 1995. We aggregated these cars into 10,000 mile odometer bins, and limited our analysis to those model year/mileage points with at least 600 cars.

Figure 7 indicates another problem with the data: we found many early model year cars with low mileages (under 100,000). One would expect the peaks of each model year distribution to move to the right with older model years; instead, the peaks for the older model years are all near 100,000 miles. We suspect that a large number of cars from these model years have 5-digit odometers, and that test technicians could not determine when an odometer had rolled over. Figures 8 through 10 show the effect the 5-digit odometers have on average emissions. For earlier model years, the average emissions of low mileage cars are higher than emissions of high mileage cars. To account for the 5-digit odometer problem, we disregard pre-87 cars, and some early model year/low mileage points on the curves.

Figure 6. Number of MY83-94 Cars by Odometer, 1995 AZ IM240

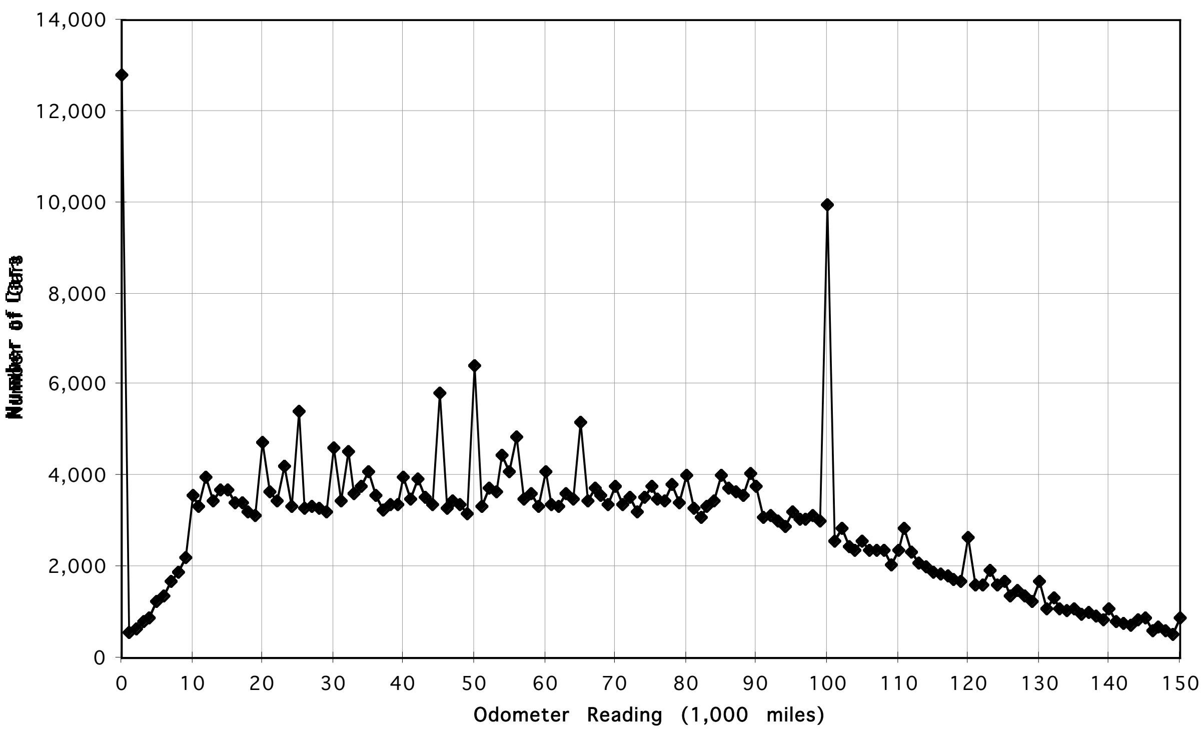


Figure 7. Number of Cars, by MY and Mileage, 1995 AZ IM240

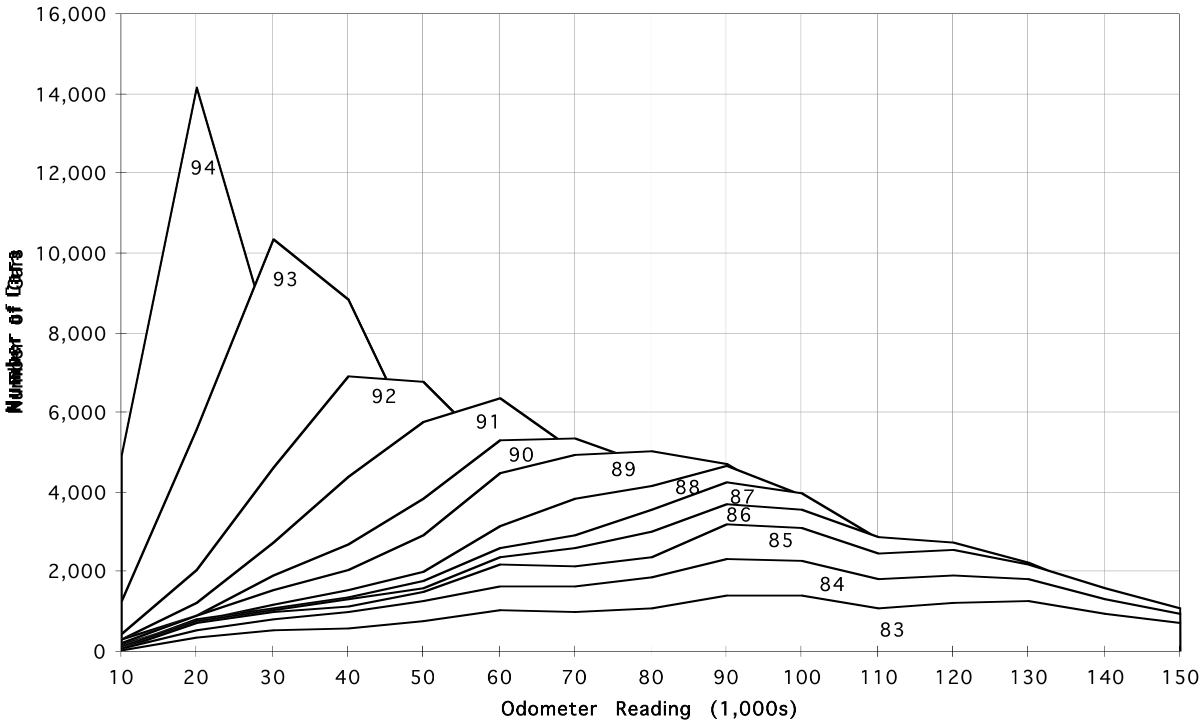


Figure 8. Average Car HC Emissions by MY and Mileage
1995 AZ IM240 (n=350,000, each point >600 cars)

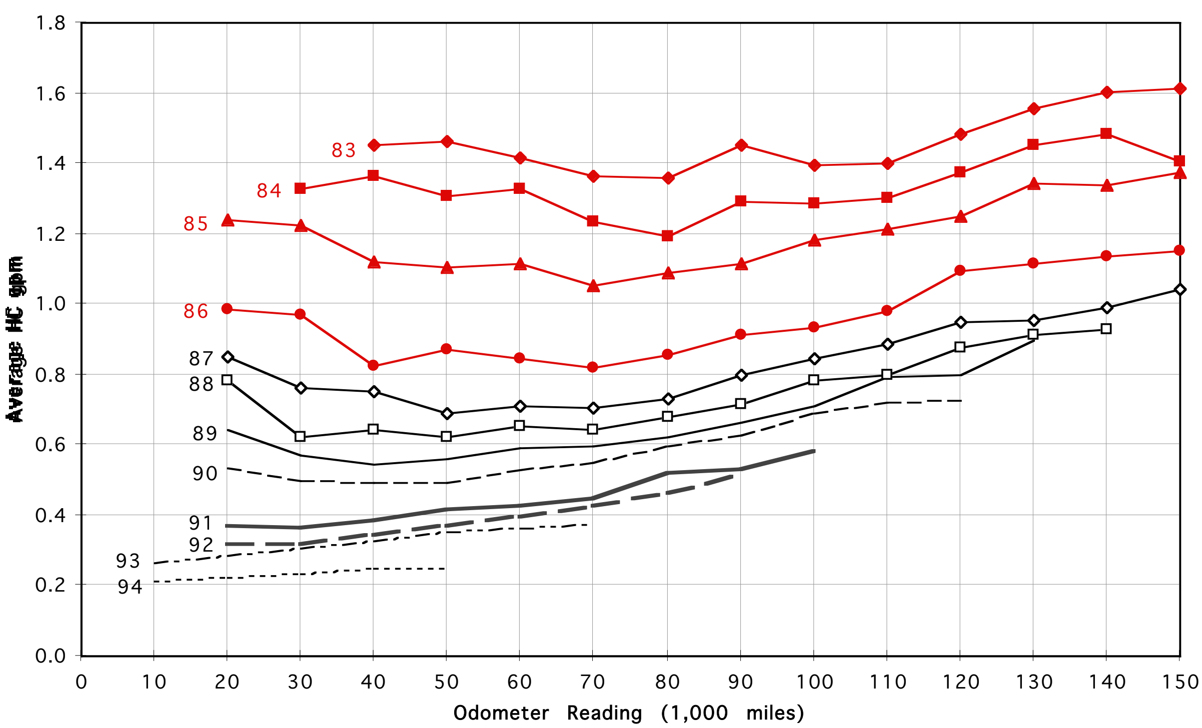


Figure 9. Average Car CO Emissions by MY and Mileage
1995 AZ IM240 (n=350,000, each point >600 cars)

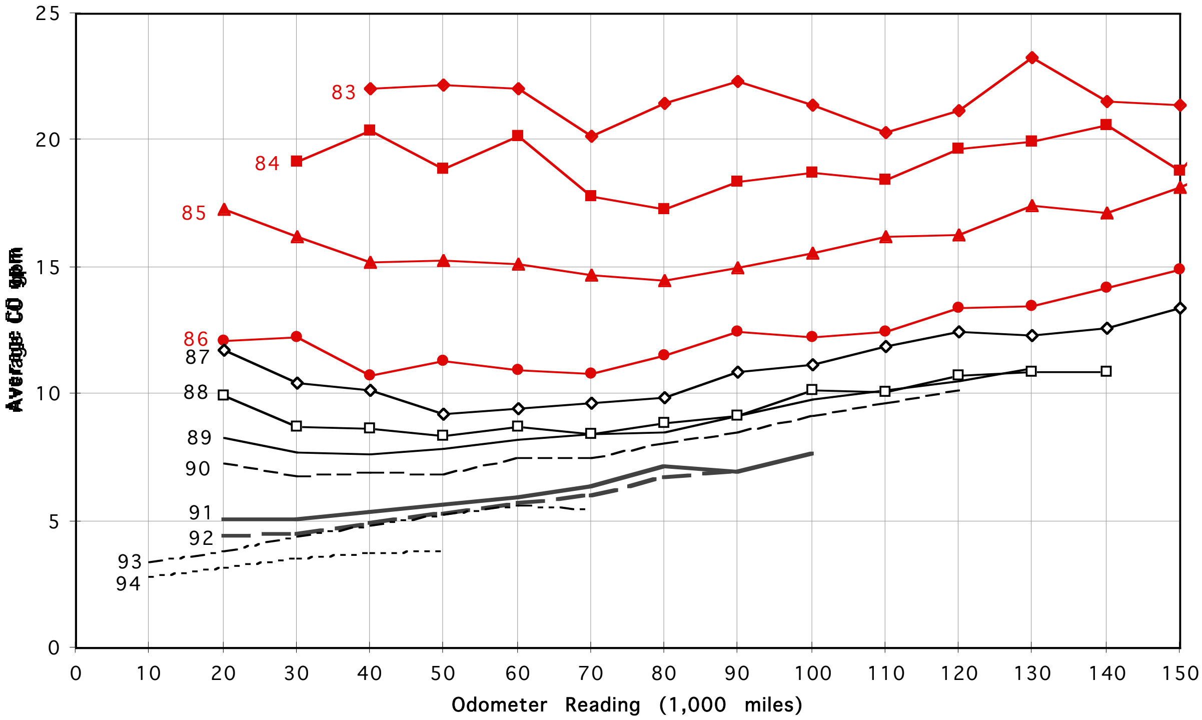
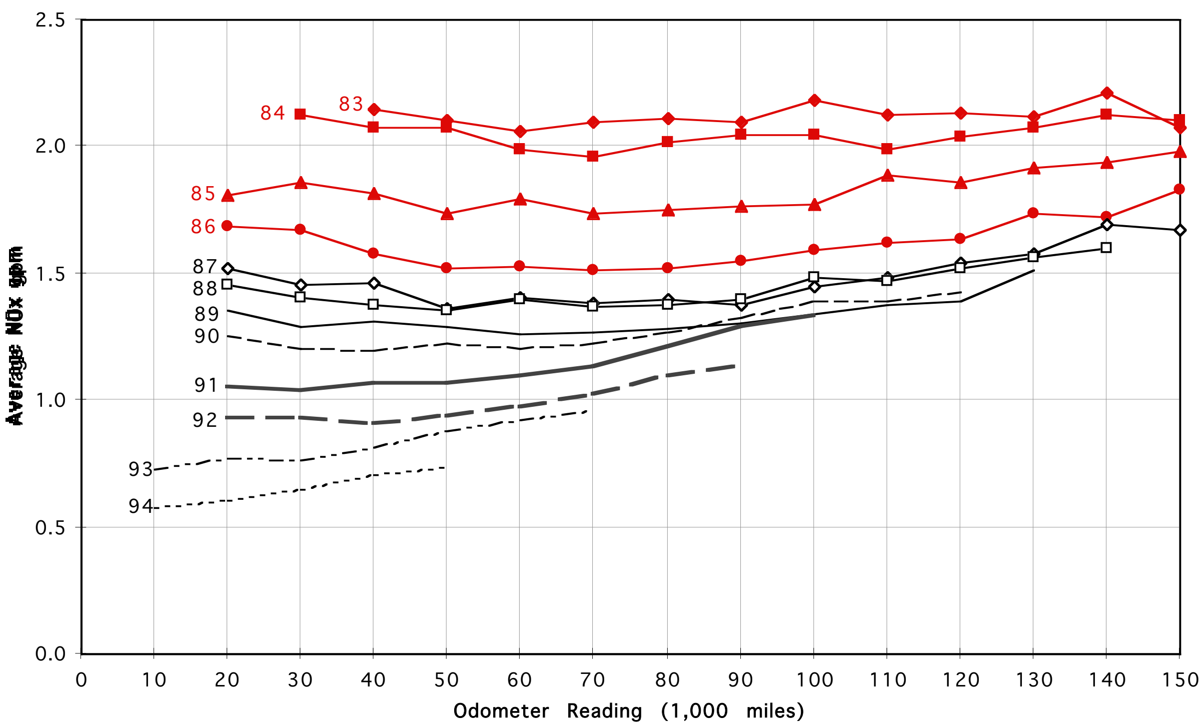


Figure 10. Average Car NOx Emissions by MY and Mileage
1995 AZ IM240 (n=350,000, each point >600 cars)



Results

Figures 11 through 13 show the emissions deterioration curves by model year and mileage. We make three observations about these curves: that emissions increase with increasing mileage; that emissions decrease with increasing model year; and that the emissions reduction by model year is greater in certain years.

Our first observation, that emissions increase with increasing mileage, is not surprising. However, we find that there is some evidence that emissions deterioration rates are lower for newer model years; for instance, newer model years have slightly shallower slopes for HC, while MY94 has a slightly shallower slope for CO. Slopes for NOx curves appear similar across all model years. All the deterioration curves by model year appear fairly linear, with no obvious changes in slope with increasing mileage, as MOBILE5 predicts. (Bows at early model year/low mileage appear to be caused by inaccurate odometer readings due to 5-digit odometers)

Our second observation, that emissions decrease with increasing model year, is somewhat surprising, in that the emissions decreases are rather large. For example, low mileage (50,000 mile) MY90 cars emit 2 to 3 times that of low mileage (10,000 mile) MY94 cars. An obvious explanation would be technological improvement that results in reduced emissions; however, if certification standards did not change between MY87 and MY93, can technological improvement by itself explain these large reductions? We speculate that aging of vehicles independent of mileage may also have an effect on emissions deterioration.

To test this, we compared 7 months of data from 1995 and 1996 to see if one year of aging has a noticeable effect on average emissions. We were surprised to find that, holding model year and mileage constant, HC and NOx emissions are consistently lower in 1996 than in 1995 (Figures 14 and 15); that is, one year older cars pollute less. CO emissions in these two years are about the same. This result is statistically significant for some model years, as shown in the figures. Since Arizona has a biennial inspection program, this reduction in emissions cannot be attributed to the effectiveness of I/M (the cars tested in 1995 are not the same as those tested in 1996). We speculate that changes in the IM240 testing, or perhaps a different mix of models in the fleets tested in 1995 and 1996, explain this difference. This surprising result does not support our claim that vehicle age affects emissions. On the other hand, one year may not be enough time to see an aging effect; clearly this issue needs more study.

Our third observation is that the trend in deterioration by model year is greater in certain model years. Figures 11 through 13 show fairly large decreases in emissions by mileage in MY91 for HC and CO (and in MY92 for NOx), and in MY94 for all pollutants. This result is likely caused by technological differences between model years. For instance over the 1980s, manufacturers replaced carburetors with fuel injection; by 1990 virtually all vehicles have fuel injection technology. The shift between the average emissions curves by model year may be due to a dramatic shift in the fraction of vehicles using fuel injection technology between MY90 and MY91. The lower MY94 curve may be due to implementation of the stricter Tier 1 emission standards; half of each manufacturer's MY94 car fleet had to meet those standards. A second possible explanation is that the stricter IM240 cutpoints for MY91 and later cars are somehow affecting the emission averages by model year.

Factors Driving Deterioraton

Finally, we tried to determine what factors are driving overall emissions deterioration. There are basically three possibilities: increasing average emissions from clean cars (cars passing the IM240); increasing average emissions from dirty cars (cars failing the IM240); and the fraction of dirty cars (as measured by the IM240 failure rate). The following results are sensitive to the cutpoints we used. We used the final IM240 cutpoints from the Arizona program (and recommended by EPA), shown in Table 1 (Arizona did not adopt these cutpoints in 1997 as planned, in part because inconsistent preconditioning would have resulted in unacceptably high false failure rates). The table shows the ratio of final IM240 cutpoints to the standards used for the Federal Test Procedure (FTP) for new car certification. Final Arizona IM240 cutpoint for CO is high relative to those for HC and NOx, meaning that fewer cars fail for CO than for HC or NOx.

Table 1. Federal Test Procedure and Arizona Final IM240 Cutpoints

Pollutant	FTP composite standards (gpm)	Final IM240 cutpoints (gpm)	Ratio of IM240 to FTP
HC	0.41	0.8	2:1
CO	3.40	15.0	4:1
NOx	1.00	2.0	2:1

Figures 16 through 21 show data from MY93 cars, representing cars at low mileages, in the left panel of each figure, and MY88 cars, representing cars at high mileages, in the right panel of each figure. We use these two model years since combined they provide emissions data on cars with odometers from 10 to 140,000 miles. The figures include the percentage deterioration rate for each curve, as well as the percent difference between each MY88 and MY93 curve. The right panel represents the worst case future for MY93 cars when they accumulate mileage; that is, that they will behave like MY88 cars at high mileage.

Figures 16 through 18 show average emissions from all cars (filled squares) and from passing cars (open squares). The difference between the two curves is the effect of failing cars on the overall emissions deterioration. We find that passing cars account for most of overall deterioration in MY93/low mileage cars; the overall deterioration rate is only slightly higher than the passing car deterioration rate. We also find that failing cars account for over half of HC, and about half of CO and NOx, deterioration in MY88/high mileage cars. The difference between the MY88 and MY93 curves at 70,000 miles represents the effect of improved technology, or vehicle age, on emissions. Technological improvement has a much bigger effect on overall emissions than on passing car emissions

Figures 19 through 21 examine the 2 components of emissions from failing cars: average emissions on the left scale, and failure rates on the right scale. Average emissions (indicated by filled squares) from failing cars increase slightly with increasing mileage; failure rates (open diamonds) show greater increases with increasing mileage. But failure rates increase dramatically by model year: MY88 failure rates at 70,000 miles are 2 to 3 times those of MY93 cars. Therefore, failure rate, rather than average emissions, is mostly responsible for the increase in emissions from failing cars.

Figure 11. Average Car HC Emissions by MY and Mileage
1995 AZ IM240 (n=220,000)

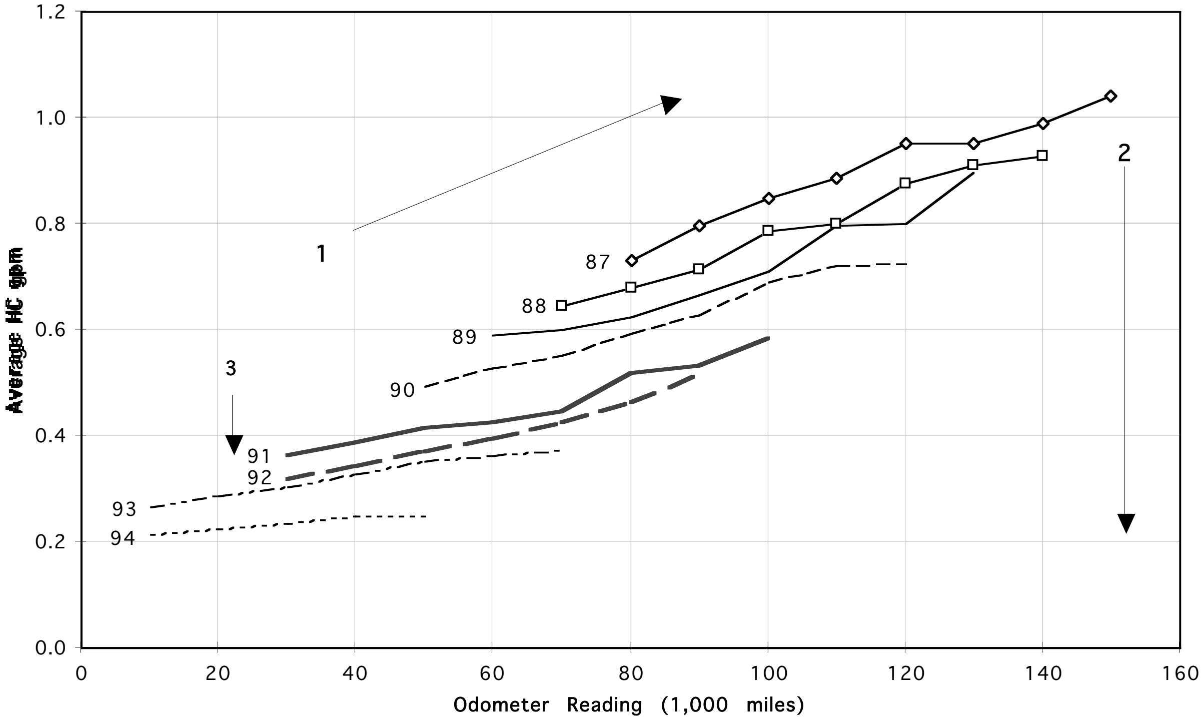


Figure 12. Average Car CO Emissions by MY and Mileage
1995 AZ IM240 (n=220,000)

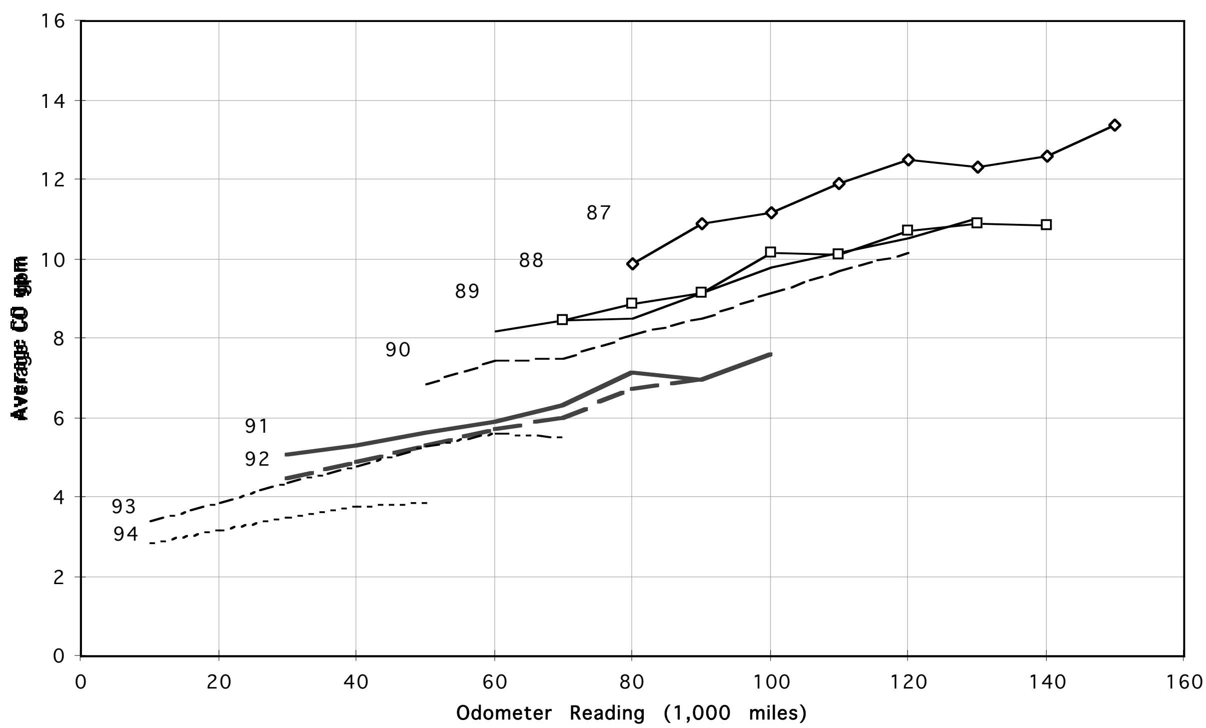


Figure 13. Average Car NOx Emissions by MY and Mileage
1995 AZ IM240 (n=220,000)

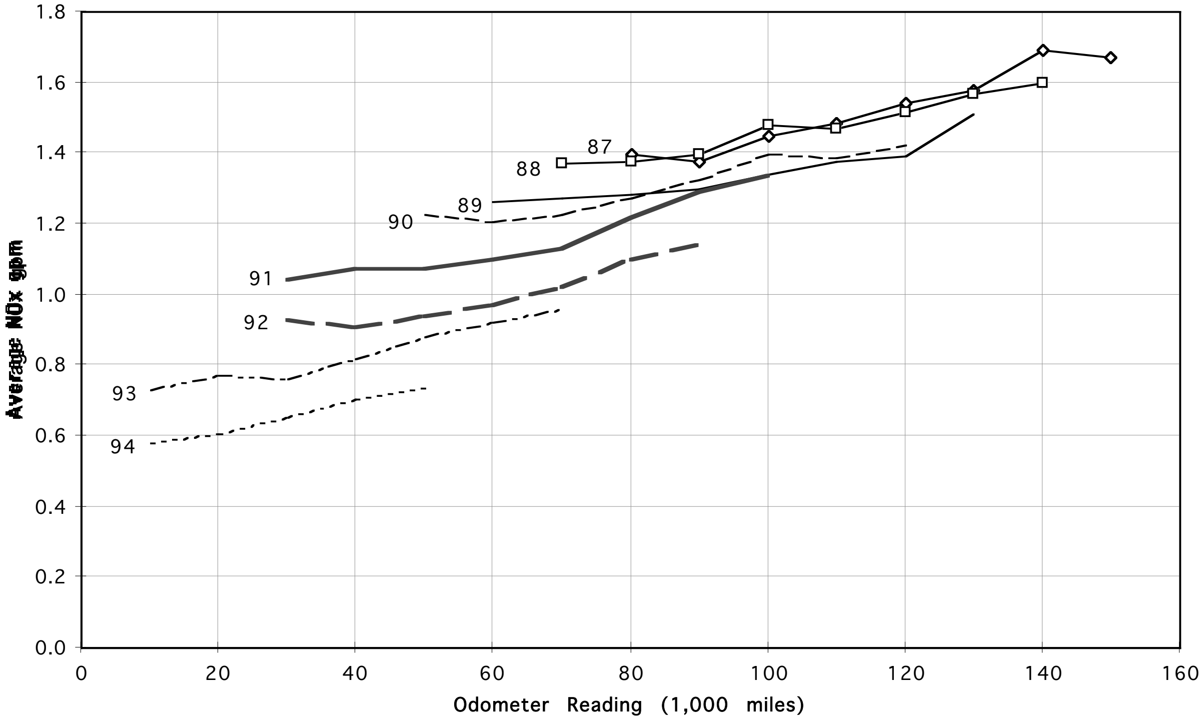


Figure 14. Average Car HC Emissions by MY and Mileage,
1995 vs. 1996 AZ IM240 (n=40,000)

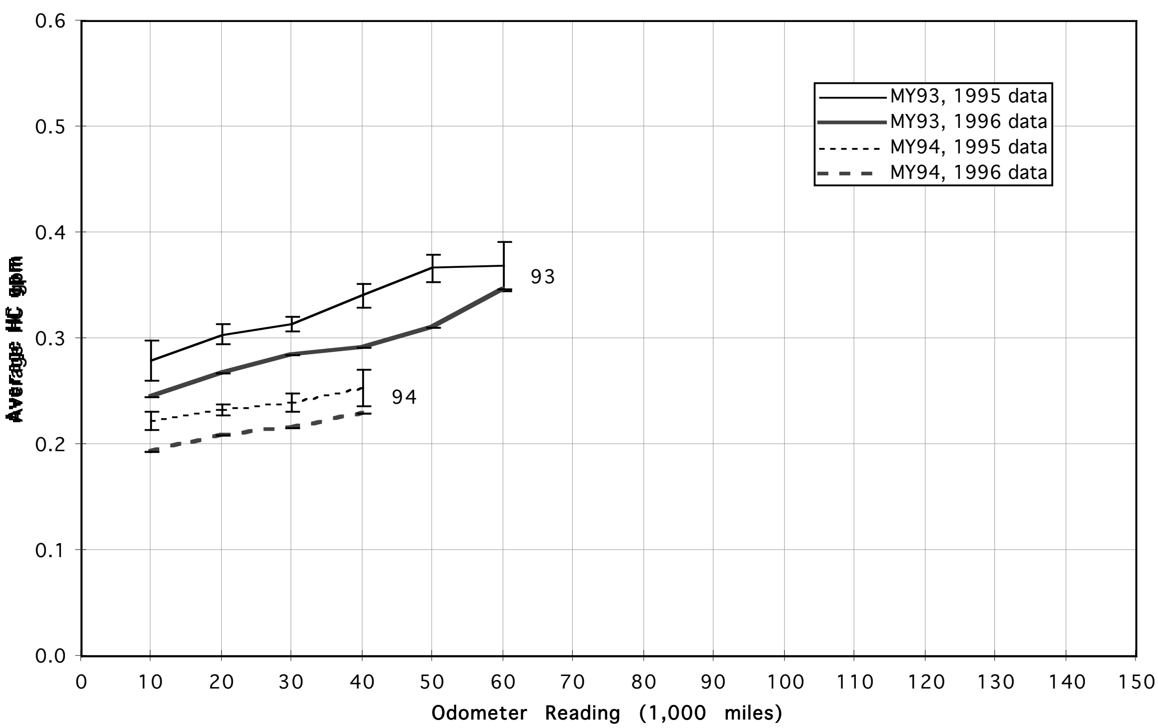


Figure 15. Average Car NOx Emissions by MY and Mileage, 1995 vs. 1996 AZ IM240 (n=40,000)

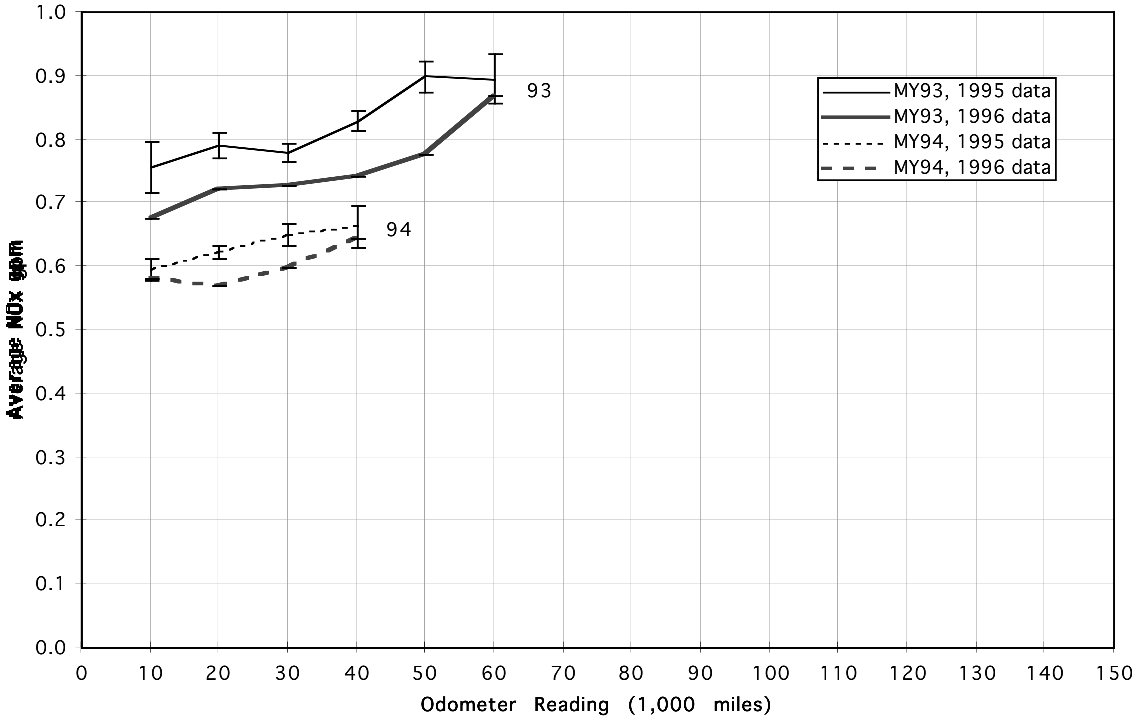


Figure 16. Average HC Emissions by Mileage, for Passing and All Cars, 1995 AZ IM240 (n=57,000)

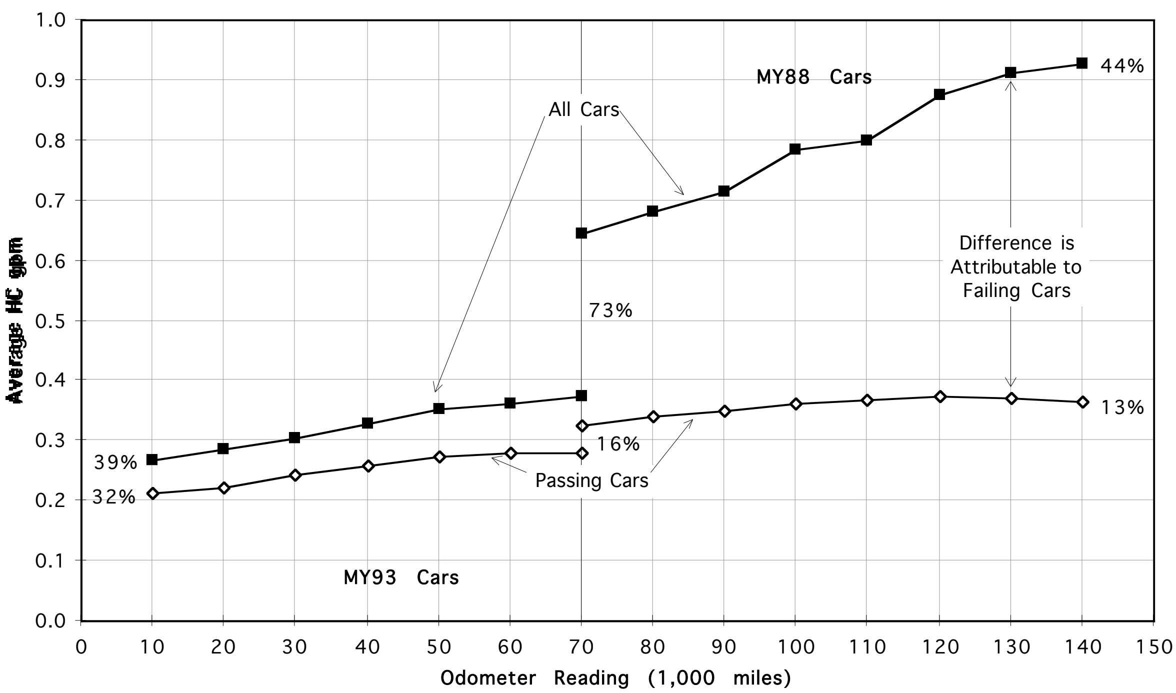


Figure 17. Average CO Emissions by Mileage, for Passing and All Cars, 1995 AZ IM240 (n=57,000)

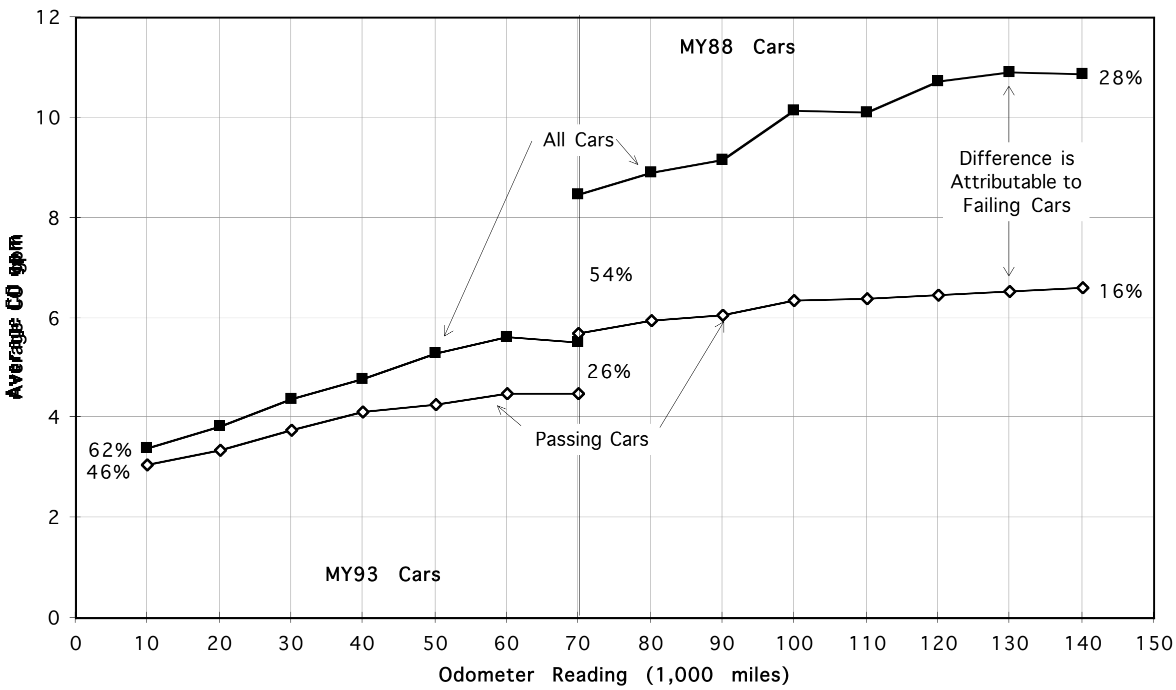


Figure 18. Average NOx Emissions by Mileage for Passing, All Cars 1995 AZ IM240 (n=57,000)

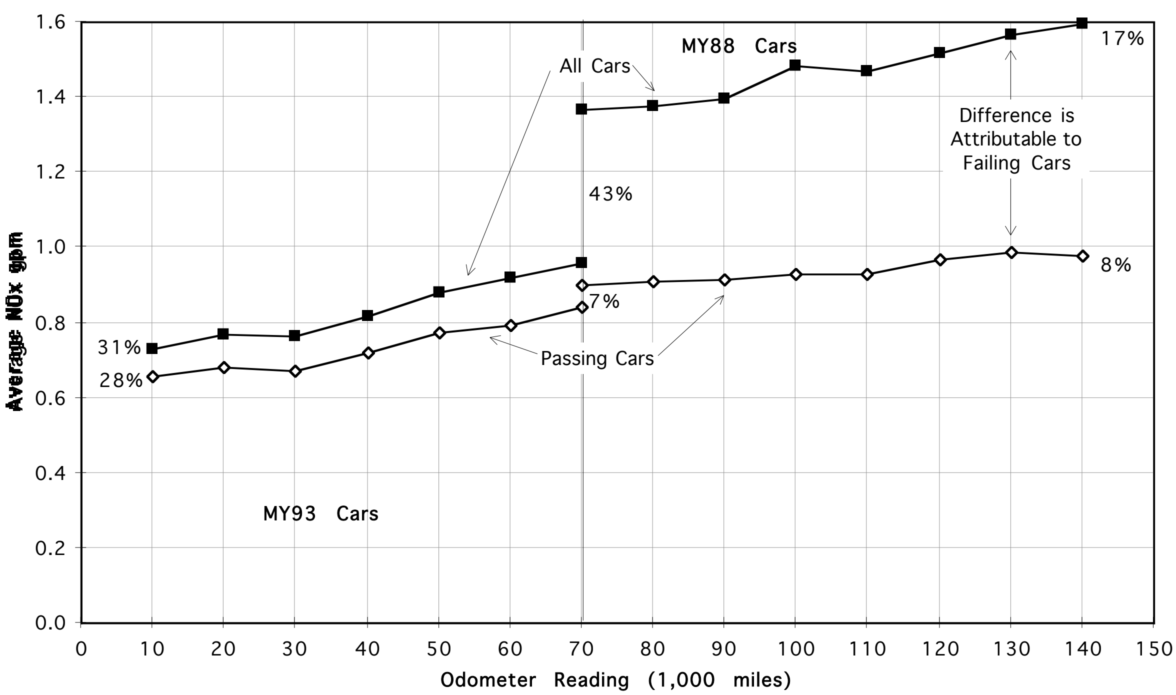


Figure 19. Components of Failing Car Contribution to Average HC
by Mileage, 1995 AZ IM240 (n=57,000)

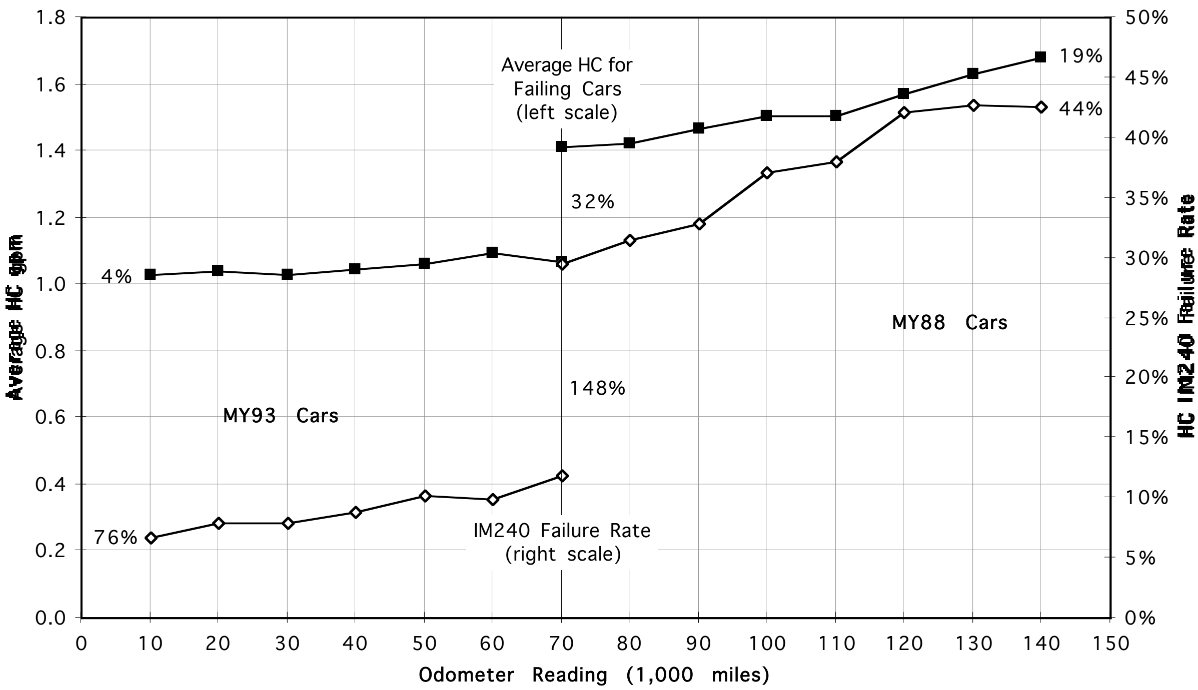


Figure 20. Components of Failing Car Contribution to Average CO
by Mileage, 1995 AZ IM240 (n=57,000)

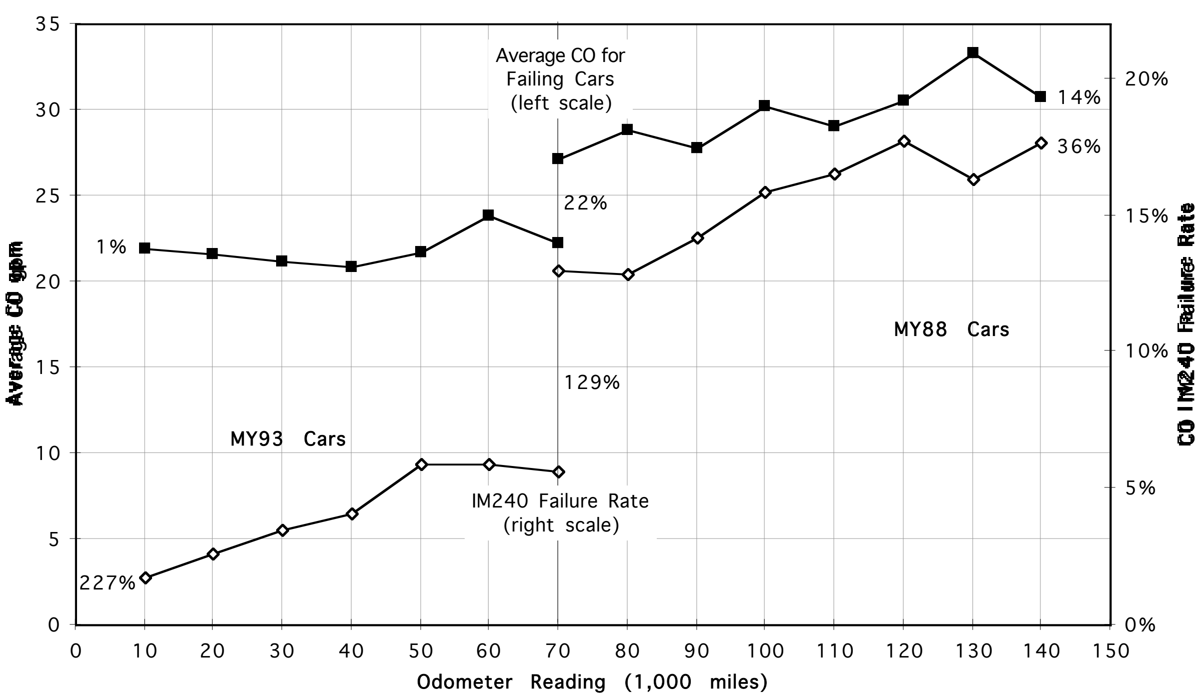
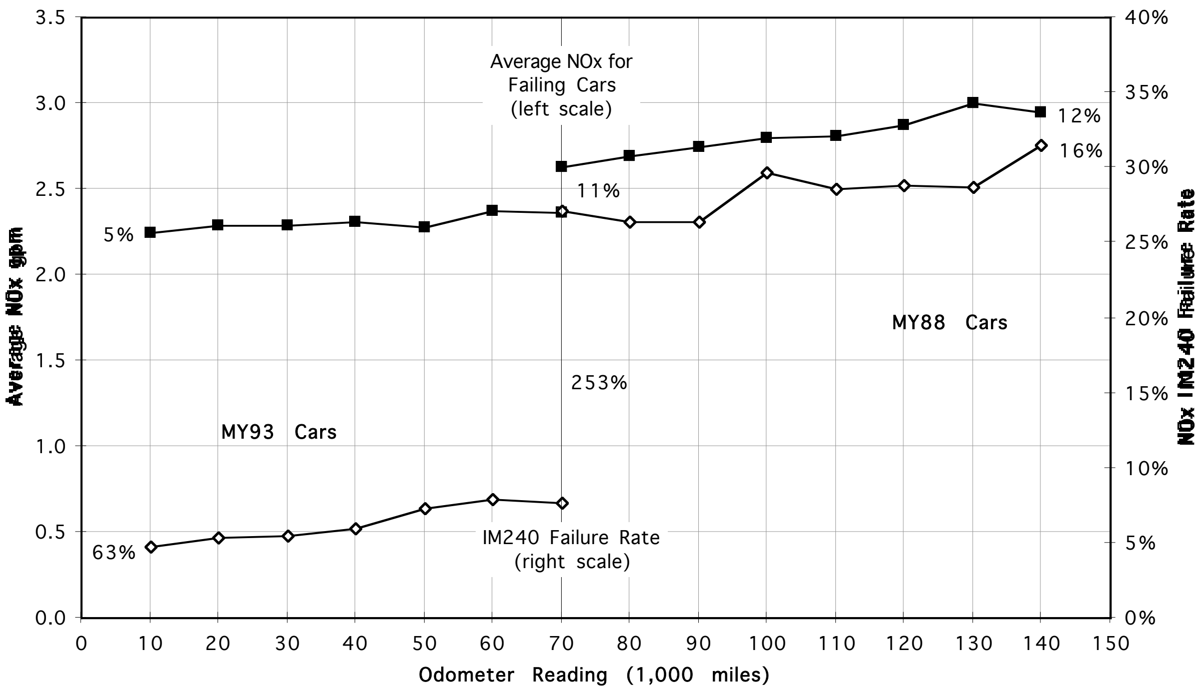


Figure 21. Components of Failing Car Contribution to Average NOx by Mileage, 1995 AZ IM240 (n=57,000)



Summary

We examined 1995 IM240 data from Arizona to determine the factors that are driving emissions deterioration as in-use vehicles age. For each model year, emissions increase linearly with increasing mileage. Emission rate curves do not kink upward with increasing mileage, as the MOBILE5 model predicts. HC emissions deterioration rates by mileage (i.e. the slopes of the curves) appear to be decreasing slightly with newer vehicle technologies, whereas the slopes of the CO and NOx deterioration curves appear fairly constant over model years.

For each mileage bin, there are large decreases in emissions with increasing model year. Most of this improvement in emissions in new vehicles is likely due to technological improvements. We suspect that vehicle age independent of mileage may be a contributing factor. When we test whether vehicle age affects emissions, we find that older cars of the same model year and mileage have higher emissions than younger cars. There are relatively large reductions in emissions in MY91 and MY94 cars. Possible causes are technological differences between model years, such as a large shift to multi-port fuel injection in MY91 and introduction of Tier 1 standards in MY94. Another possible cause is the stricter IM240 emission cutpoints used for MY91 and newer cars.

We use average emissions from older (i.e. MY87) vehicles to simulate emissions at high mileage. Using these data, it appears that vehicle failure rate, rather than increasing average

emissions from passing cars or increasing average emissions from failing cars, has the biggest impact on overall emissions deterioration as vehicles accumulate mileage.

We plan to continue analyzing I/M data from Arizona and other states in several ways to learn more about emissions deterioration. Most importantly, we would like several test years of data to better examine the effect of aging independent of mileage on emissions. We also plan to:

- test the sensitivity of our results to the cutpoints we use (particularly for CO);
- identify early model year cars with legitimate odometer readings, by model year and manufacturer or model, in order to include more early model year/low mileage points in our analysis;
- determine why one year older cars (tested in 1996) had lower emissions than the cars tested in 1995; and
- study deterioration rates by type of fuel system (carbureted vs. throttle body vs. multi-point fuel injection) and by vehicle type (car vs. light duty truck).

References

Heirigs, Philip L. and Jay Gordon. 1996. *Preconditioning Effects on I/M Test Results Using IM240 and ASM Procedures*. SAE Technical Paper Series 962091. October.

981414

Characterization of Recent-Model High-Emitting Automobiles

Tom Wenzel
Lawrence Berkeley National Laboratory

Marc Ross
University of Michigan

ABSTRACT

In-use vehicles which are high emitters make a large contribution to the emissions inventory. It is not known, however, whether high-emitting vehicles share common emissions characteristics. We study this by first examining laboratory measurements of second-by-second engine-out and tailpipe emissions from a small number of MY90-97 high-emitting vehicles. We distinguish high-emitter types by the behavior of six ratios in low- and moderate-power driving: the engine-out emissions indices (engine-out pollutant to fuel-rate ratios) and the catalyst pass fractions (tailpipe to engine-out ratios) for CO, HC, and NOx. Four general types of high emitter are observed: 1) fuel-air ratio excessively lean, 2) fuel-air ratio excessively rich, 3) partial combustion such as misfire, and 4) severe deterioration in catalyst performance in vehicles where malfunctions of Types 1, 2 or 3 are not predominant. We also find that these behaviors may be chronic, or may only occur transiently. The second step is to determine the prevalence of the four different types of high emitter in the on-road fleet. For this we analyze IM240 tailpipe emissions from a large sample of cars measured in the Arizona inspection and maintenance program. We find that all four types of failure are observed with roughly comparable probabilities.

INTRODUCTION

Several independent analyses have found that about half of the on-road emissions by automobiles may be from the small fraction of vehicles that are high emitters [1-4]. Although there are many potential technical causes of failed or malfunctioning emissions controls, there has been relatively little study of the distribution of these technical causes in the fleet of in-use vehicles [5-7]. Probably the most useful work is a comprehensive analysis of several datasets on the effectiveness of repairing specific components, which identifies components most likely to fail [8,9].

In the nature of investigations of high-emitters, the emphasis has been on carbureted vehicles and early-model fuel-injected vehicles. In the present analysis, we focus on newer model years, presenting information on model-year 1990 and later vehicles with sophisticated computer-controlled fuel-injected engines.

First, we identify the types of high emitters in hot-stabilized operation, and draw rough conclusions about the physical mechanisms underlying each, based on detailed second-by-second testing of engine-out and tailpipe emissions on a sample of in-use vehicles at the University of California, Riverside. In particular, we distinguish high-emitter types by the behavior of six ratios in low- and moderate-power driving: the engine-out emissions indices (engine-out pollutant to fuel-rate ratios) and the catalyst pass fractions (tailpipe to engine-out ratios) for CO, HC, and NOx. Thus our determinations of the causes of high emissions are based on detailed comparisons of fuel rate, and engine-out and tailpipe emissions, rather than on mechanical inspection or any subsequent emissions reductions due to component repairs and/or replacements.

Second, we estimate the frequency of occurrence of each type of malfunction in the in-use fleet, based on analysis of results from the inspection and maintenance (I/M) program in Arizona. The distribution of three-pollutant "profiles" in the I/M data enables estimation of the on-road probabilities for each type of high emitter observed in the laboratory measurements made at UC Riverside.

HIGH EMITTER TYPES IN THE NCHRP DATA

A major emissions measurement has been recently completed at the College of Engineering Center for Environmental Research and Technology (CE-CERT) at the University of California at Riverside, funded by the National Cooperative Highway Research Program (NCHRP) of the Transportation Research Board. The primary purpose is to develop a modal, or driving-dependent, emissions model [10, 11]. Both engine-out and tailpipe emissions of some 300 vehicles have been measured second-by-second on three driving cycles, including the Federal Test Procedure (FTP) cycle and a modal cycle developed at CE-CERT for modeling purposes (the Modal Emission Cycle, or MEC). The vehicles have been recruited for emission model development; i.e., in accordance with their relative contribution to the emission inventory, rather than according to their frequency on the road. The emissions of roughly fifteen MY1990 and later high-emitting cars and a similar number of high-emitting light-duty trucks are among the vehicles recruited and measured. (The number of high emitters depends on the cutpoints used to define high emissions.) Sixteen of these high-emitters are analyzed in this paper. The NCHRP project is the first to specifically recruit high-emitters for such second-by-second measurements of both tailpipe and engine-out emissions.

The recruitment of high-emitting vehicles of MY1990 and later is difficult because the fraction of such vehicles in the fleet is low (at least at current vehicle ages). In the NCHRP project vehicles suspected of being high emitters were specifically recruited in a non-random fashion, so the overall frequency of high emitters, and the frequency by type of failure, in the on-road fleet is not known from these data. (It should be clear that the identification of one or two vehicles of a particular model as high emitters in this project has no statistical significance.)

To address the issue of real-world frequency of the high emitters, we categorize the several types of high emitters measured in the project according to their emissions characteristics, and make a correspondence between these types of high emitter and the distribution of high emitters with similar tailpipe-emission profiles observed in Arizona's on-going I/M program. The Arizona program covers essentially all light-duty vehicles in the Phoenix area (although the number of high emitters may be underestimated because there is a tendency for people to not register their vehicles, or register them elsewhere, if they think that they won't pass the I/M test [12]). We thus determine weights to assign to the NCHRP high-emitter types which may reasonably reflect the representation of those kinds of high emitters on the road.

The characterization of the NCHRP high emitters might be done using simulation-model parameter fits to the measurements, or simply from bag data. But emissions in distinct driving modes will be used here because it is a simple approach which reveals aspects of the physical mechanisms of emissions control system (ECS) failure. (Note that careful inspection of the tested vehicles by a professional mechanic was not a part of the NCHRP project.)

We focus our study on vehicles which are high emitters in low- to moderate-power driving. An example of what we call moderate power is a 50 mph cruise on a level road without unusual load, but with throttle fluctuations. Such a power level requires a fuel rate of about 0.7 grams per second for small sedans, and about twice that for large sedans and most light trucks. This power level is characteristic of the IM240 driving cycle used in the Arizona I/M program and the 505-second cycle used for bags 1 and 3 of the Federal Test Procedure (FTP), as shown in Table 1. Such moderate power modes are also found in the MEC. The maximum fuel rates achieved in throttle fluctuations during the MEC are also shown in parentheses and are seen to be less than the maxima in the regulatory cycles.

Table 1. Modes of the MEC Considered

Mode	**Avg speed (mph)**	**Avg (Max) Fuel Rate (g/s)** small sedan	**Avg (Max) Fuel Rate (g/s)** large sedan
MEC:			
low power	20, 35	0.4 (0.7)	0.6 (1.2)
mod. power	50	0.7 (1.1)	1.3 (2.0)
IM240		0.7 (2.1)	1.2 (3.5)
FTP Bag 2		0.4 (1.3)	0.8 (2.2)
FTP Bag 3		0.6 (2.1)	1.0 (3.5)

We will compare emission rates in the MEC, and Arizona IM240 measurements (as well as referring to analyses of earlier FTP measurements). As seen in Table 1, the average and maximum power levels in FTP bag 2 are substantially less than in the IM240 cycle, while bag 3 and IM240 have similar power levels. On the other hand, bag 3 starts after a 10 minute soak which modestly increases CO and HC totals for the bag. The IM240 is supposed to begin with the vehicle hot, but there is evidence that in practice vehicles often may have cooled off somewhat or the engine block may not have been fully warmed up [13]. Power levels and vehicle conditioning in the selected modes of the MEC are most comparable to those of FTP bag 3 and the IM240 cycle.

EMISSIONS BEHAVIOR IN CLOSED-LOOP AND COMMAND ENRICHMENT

Accurate control of the fuel-air ratio in closed-loop operation is critical to effective emissions control. It is likely that most high emitters among MY1990 and later vehicles are caused or created by some form of fuel-air ratio control problem.

In closed-loop operation with a three-way catalyst, the electronic control module manages the injection of fuel so as to essentially maintain stoichiometry (the optimum ratio of air to fuel, about 14.7:1) to maintain combustion while minimizing emissions. In vehicles with three-way catalysts, the ratio is made to swing back and forth between slightly rich and slightly lean, at about 1 Hz or faster, in order to automatically adjust the oxygen level on catalyst surfaces so that exhaust CO and HC are oxidized while NO is simultaneously reduced. The time dependence of the fuel-air ratio in a typical properly-functioning vehicle is schematically shown in Figure 1. As shown, for proper operation the fuel-air ratio oscillates around stoichiometric:

$$\tfrac{1}{2}\langle f\rangle - 1\tfrac{1}{2} < Df \qquad (1)$$

Here, f is the fuel-air ratio compared to its stoichiometric value. In fact, eq(1) should hold with substantial overlap. For many vehicles with malfunctioning ECS the fuel-air management isn't working properly, so this inequality doesn't hold, even at moderate power. In these conditions, the vehicle is likely to be a high emitter.

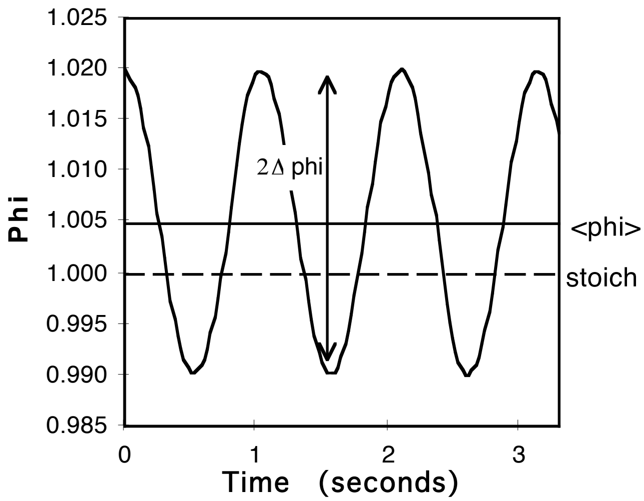


Figure 1. Illustrative Example of Oscillations in Fuel-Air Ratio in Closed-Loop Operation

In Table 2, six emissions ratios measured in the NCHRP project are shown with typical values that have been observed for modern *properly-functioning* vehicles in hot-stabilized operation (specifically, MY91-93 vehicles tested by manufacturers as part of the FTP Revision Project [14,15]). We distinguish three fuel-air ratio regions: *stoichiometric*, where eq(1) is satisfied; *rich*, where f > 1 beyond that described by Figure 1; and *lean*, where f < 1 beyond that described by Figure 1.

Table 2. Average Emission Ratios for Low-Emitting Vehicles, Stoichiometric and Rich Operations

	Operating Range	
Variable	**Stoichiometric**	**Enrichment**
EICO	» 0.08	0.1 to ~1.0
EIHC	» 0.015	» 0.015
EINOx	£ 0.05, lower at low power	£ 0.05, declines with enrichment
CPFCO	£ 0.1	quickly ® 1.0
CPFHC	£ 0.1	gradually ® ~0.7
CPFNOx	0.02 to 0.2	quickly ® ~0.7

In stoichiometric operation one observes that:

- The CO emission index, or EICO (the ratio of mass of CO that leaves the engine to fuel input mass), varies around 0.08, from perhaps 0.02 to 0.15.
- EIHC depends somewhat on details of engine design and fuel and lubricant composition, since it comes from cylinder surfaces and crevices; but it lies between 0.01 and 0.02 in rich as well as stoichiometric operation.
- EINOx, the engine-out NOx-to-fuel mass ratio, varies with power and with EGR system. The typical maximum value observed is 0.05.
- We designate catalyst activity using catalyst pass fractions, or CPFi: the mass ratio of pollutant i output from the catalyst to pollutant i input to it (i.e. the tailpipe to engine-out ratio). The three catalyst pass fractions vary considerably from one vehicle model to the next and with the details of operation.

In high-power operations, most vehicles command fuel enrichment; i.e. the fuel-air control system goes open loop and f is commanded to be in a range roughly 1.05 to 1.20 (i.e. 5 to 20 percent rich). Since command enrichment results in massive increases in tailpipe CO emissions and some increase in HC, and will, moreover, be coming under regulation with the Supplemental FTP, manufacturers have begun to reduce the use of this technique.

The emissions ratios behave in predictable ways when the fuel-air ratio goes rich (right-hand column, Table 2):

- EICO increases strongly with enrichment (as shown by eq(2), below); CPFCO is sensitive to even slight enrichment and increases rapidly toward 1.0 with increasing enrichment.
- EIHC is essentially independent of enrichment as such because at the high cylinder temperatures excess fuel is converted to CO and H_2; however, it increases due to other kinds of incomplete combustion, such as from cylinder misfire. CPFHC increases slowly with increasing enrichment.
- EINOx is moderately suppressed by the cooling effect of enrichment; CPFNOx may be reduced with slight enrichment, but increases rapidly with stronger enrichment in most modern vehicles (although it does decline in a few models).

In decelerations during closed-loop operation the fuel-air ratio often goes lean, often very lean in major decelerations. Lean excursions are normal, although large engine-out HC puffs may occur. If catalyst performance has deteriorated, then tailpipe HC puffs associated with these lean excursions can be substantial [16].

FUEL-AIR RATIO DATA

As suggested by Figure 1, f (the fuel-air ratio relative to stoichiometric) would need to be known to much better than 2% accuracy to be useful for our purposes here. Fuel-air ratios based on emission measurements and chemistry are not accurate enough for this purpose. For this reason we use the emissions ratios listed in Table 2 as indicators of improper fuel management.

As an alternative to calculating f from tailpipe measurements and chemistry, one can estimate it from a linear formula for EICO:

EICO » 0.08 + 3.6(1 - 1/f), or

f = 1 + (EICO - 0.08)/(3.5 - EICO) (2)

It is likely that f calculated using eq(2) is not grossly in error. Eq(2) is not however useful in lean conditions.

DEFINITION OF HIGH AND LOW EMITTERS

For this paper, we define high emitters in the NCHRP project as vehicles which exceed FTP bag 3 emissions cutpoints in grams per mile (gpm); the selected cutpoints are shown in

Table 4 below. With the chosen cutpoints, high emitters exceed the emissions of typical properly-functioning MY 1990-1993 vehicles by more than a factor of about 2.5. These are rather tight cutpoints for "high emitters"; we choose them because MY90 and later high emitters proved hard to recruit for testing.

For our analysis we also need cutpoints below which we consider a vehicle to be a low emitter. For this purpose we examine three sets of measurements, as summarized for cars in Table 3. The measurements are: 1) NCHRP, for MYs 90-93 measured in 1996-97 (mostly California cars). We calculate average emissions for properly-functioning cars by excluding the 10% highest emitters. 2) FTP Revision Project measurements on new MY91-94 49-state vehicles with 50,000 mile laboratory-aged catalysts [17]. 3) American Automobile Manufacturers Association in-use survey from which we select MY 1991-92 cars with odometer readings from 40,000 to 60,000 miles, measured in 1995-96 [18]. Again, we take the average emissions of the 90% cleanest cars (sorted for each pollutant separately).

Table 3. Emissions from Properly-Functioning Cars at 50,000 miles in Three Studies: FTP Bag 3 (gpm)

dataset	MYs	n[a]	CO	HC	NOx
NCHRP	1990-93	24	2.7	0.22	0.35
FTP-RP	1991-94	23	1.5	0.16	0.33
AAMA in-use	1991-92	57	2.5	0.21	0.22

a) number of vehicles measured in the subset considered. See text for definition of each subset.

The low cutpoints adopted are shown in Table 4. We regard these low cutpoints to be representative of properly-functioning in-use vehicles at 50,000 miles and age 4 to 5 years. Roughly two-thirds of properly-functioning vehicles will emit less than the low-emitter cutpoints chosen.

Table 4. Cutpoints for High and Low Emitting Vehicles in the NCHRP Project: FTP Bag 3 (gpm)

	CO	HC	NOx
Low Emitters			
cars	3	0.2	0.4
trucks	4	0.3	0.7
High Emitters			
cars	6	0.5	1.0
trucks	10	0.8	1.5

HIGH-EMITTER TYPES

Below we consider the four types of high emitters observed in NCHRP project measurements

Type 1. Operates Lean at Moderate Power

In the first type of high emitter, the fuel-air ratio is chronically lean or goes lean in transient operation calling for moderate-power. An average 2% or more lean is likely to saturate the catalyst with oxygen. The examples from the NCHRP data are vehicles 103 (1993 Sundance), 202 (1997 Windstar), and 295 (1990 Astro).

The characteristics of the six ratios for vehicle 202 at low and moderate power are shown in Table 5. The effect on the CPFs is striking, while that on the engine-out emissions is slight. While vehicle 202 operates consistently lean, vehicle 103 goes lean in moderate-power transients (i.e. with throttle fluctuation). Vehicle 295 also goes lean during transients, and shows considerable catalyst deterioration as well.

Table 5. Average Emission Ratios at Moderate Power for Type 1 (Vehicle 202)

Variable	Range, Comment
EICO	» 0.08 or less, normal
EIHC	» 0.02, normal
EINOx	£ 0.1, slightly > normal
CPFCO	» 0.01, almost zero, < normal
CPFHC	» 0.01, almost zero, < normal
CPFNOx	roughly 0.5 to 1.0, much > normal

The behavior of a high NOx emitter over a portion of the MEC (Figure 2a) is compared with that of a normal NOx emitter (Figure 2b). The tendency of vehicle 202 to run lean for long stretches is seen in Figure 2a. In driving at 50 and 65 mph, phi is frequently about 0.9, and the tailpipe NOx rate is high, reaching 0.1 or 0.2 grams per second. Vehicle 136, a normal NOx emitter, operates at stoichiometry during the cruise sections, resulting in very low tailpipe NOx levels. (The strong acceleration at approximately 110 to 120 seconds involves power beyond FTP levels which we do not consider here.)

The FTP bag 3 tailpipe emissions profile for these vehicles is shown in Table 6: very high NOx tailpipe emissions, and low CO and HC emissions, relative to emissions of clean vehicles. The profile is in the form of CO/HC/NOx levels in terms of the two cutpoints for each, with L, M and H standing for: below the low cutpoint, medium or in between, and above the high cutpoint, respectively. The low and high cutpoints for trucks are shown for comparison, from Table 4.

Table 6. FTP Bag 3 gpm Tailpipe Emissions for Type 1 Vehicles, and Truck Cutpoints

Test Vehicle	CO	HC	NOx	profile
103 (car)	1.7	0.05	1.1	LLH
202 (truck)	0.4	0.04	2.9	LLH
295 (truck)	4.0	0.90	1.8	MHH

A physical failure mechanism leading to Type 1 behavior is not so easy to pinpoint. Improper signal from the oxygen sensor or improper functioning of the electronic engine control are possibilities.

Type 2. Operates Rich at Moderate Power

In the second type of high emitter, the fuel-air ratio is chronically rich or goes rich in transient moderate-power operation. The EIHC remains normal. Under these conditions, the CO emission index and catalyst pass fraction are high, resulting in high tailpipe CO emissions. Examples from the NCHRP testing are three cars, 113 (1990 Sentra), 125 (1990 Spirit), and 136 (1993 240 SX).

The measurements on vehicle 113 at low and moderate power are summarized in Table 7. The high EICO and CPFCO occur in moderate-power transients (i.e. with throttle fluctuation). Relative to properly-functioning vehicles, EIHC is unaffected and EINOx is slightly low. The behavior of vehicle 136 is similar. Vehicle 125 shifts from stoichiometric to steady highly-enriched operation for long periods in a manner apparently unrelated to the driving. Vehicles 43 and 277 show transient enrichment, but their strong deterioration of catalyst performance leads us to categorize them as Type 4 below.

Table 7. Emission Ratios at Moderate Power for Type 2 (Vehicle 113)

Variable	Range, Comment
EICO	> 0.15, 2 or more times normal
EIHC	» 0.015, normal
EINOx	» 0.02, < normal
CPFCO	roughly 0.5 to 1.0, much > normal
CPFHC	» 0.05 to 0.2, somewhat > normal
CPFNOx	» 0.01, < normal

The behavior of a high CO emitter over a portion of the MEC (Figure 3a) is compared with that of a normal CO emitter (Figure 3b). The tendency of vehicle 136 to run somewhat rich when there are throttle variations at moderate power is shown in Figure 3a in the 60- to75-second segment, where EICO reaches levels of 0.2 to 0.3. The great sensitivity of CPFCO to these rich excursions is evident. A normal CO emitter, vehicle 103 (Figure 3b) shows much lower EICO and CPFCO in this segment of the MEC. (Again we do not focus on the strong accelerations at the beginning and end of the sequence shown.)

The FTP bag 3 tailpipe emissions profile for these vehicles is shown in Table 8: high CO, and low to medium HC and NOx, relative to emissions of clean vehicles. The low and high cutpoints for cars, from Table 4, are shown for comparison. (For car 113, the CO is taken as high although the measurement comes in slightly below the high cutpoint.)

There are many possible failure mechanisms resulting in enrichment during closed loop operation; however the mechanism here must also leave the engine-out HC emissions index in its normal range of 0.01 to 0.02. Thus there can be enrichment but not misfire. One example which meets the characteristics is a leaking exhaust line which brings in oxygen before the oxygen sensor, resulting in the sensor calling for more fuel from the injectors.

Table 8. FTP Bag 3 gpm Tailpipe Emissions for Type 2 Vehicles, and Car Cutpoints

Test Vehicle	CO	HC	NOx	profile
113 (car)	5.9	0.21	0.24	HML
125 (car)	6.4	0.34	0.57	HMM
136 (car)	6.8	0.17	0.17	HLL

Type 3. High Engine-Out Hydrocarbon Emissions Index

The third type of high emitter involves a high engine-out emission index for HC and mild enrichment, as evidenced by high EICO and CPFCO. Catalyst performance is also poor. The examples are vehicles 178 (1992 S-10 pickup), 209 (1994 Caravan), and 273 (1992 Corsica). The characteristics of vehicle 209, whose second-by-second EIHC is consistently high, are shown in Table 9.

Table 9. Emission Ratios for Type 3 (Vehicle 209)

Variable	Range, Comment
EICO	> 0.15, 2 or more times normal
EIHC	» 0.15, roughly 10 times normal
EINOx	» 0.02, < normal
CPFCO	roughly 0.5 to 1.0, much > normal
CPFHC	» 0.05 to 0.2, slightly > normal
CPFNOx	» 0.01, essentially zero

The characteristics of vehicle 178 are shown in Table 10. In this case, high EIHC is a transient effect, with puffs of HC every time the fuel-air ratio declines, even in cases where it remains rich.

Table 10. Emission Ratios for Type 3 (Vehicle 178)

Variable	Range, Comment
EICO	» 0.15, slightly over normal
EIHC	» 0.05, roughly 3 times normal
EINOx	< 0.02, < normal
CPFCO	roughly 0.5, much > normal
CPFHC	» 0.1 to 0.3, > normal
CPFNOx	» 0.5, much > normal

The behavior of a high HC emitter over a portion of the MEC (Figure 4a) is compared with that of a normal HC emitter (Figure 4b). The tendency of vehicle 178 to have HC emissions indices exceeding 0.1 at times other than major decelerations is shown in Figure 4a. The effect seems to be associated with throttle fluctuations between seconds 70 and 80 of the MEC (the relatively low EICO values at these times suggest that the increase in EIHC is not due to enrichment; an example of enrichment can be seen between seconds 40 and 45, at the end of an acceleration). Figure 4b shows that a properly-functioning engine of current technology maintains EIHC in the 0.01 to 0.02 region, except after major accelerations or decelerations. (The figure also shows small EIHC excursions above this value during transients.)

The FTP bag 3 tailpipe emissions profile for these vehicles is shown in Table 11: moderate to slightly-high tailpipe CO, very high HC, and moderate to low NOx relative to properly-functioning vehicles. The key aspect of the profile is the very high HC.

Table 11. FTP Bag 3 gpm Tailpipe Emissions for Type 3 Vehicles

Test Vehicle	CO	HC	NOx	profile
178 (truck)	4.5	1.2	0.80	MHM
209 (truck)	11.4	2.1	0.06	HHL
273 (car)	9.8	1.7	0.90	HHM

Excess EIHC is probably caused by incomplete combustion in one or more cylinders, from many physical mechanisms such as a bad spark plug or partial obstruction of an injector resulting in too little fuel injected into the cylinder. There are many possible mechanisms. Oxygen levels in the exhaust are observed to be correspondingly high (2.5 grams of excess oxygen per gram of excess engine-out fuel). Catalyst performance is also poor, and not only when EIHC is high. Perhaps the catalyst deterioration results from the history of high engine-out HC emissions.

Type 4. Poor Catalyst Performance for All Three Pollutants at Moderate Power

High tailpipe emissions of all pollutants typifies Type 4 high emitters. This type involves more than one behavior, with 1) chronically poor catalyst performance, due to burned-out or missing catalyst, or 2) transiently poor catalyst performance, e.g. a catalyst pass fraction of 0.3 or more in moderate-power driving. Type 4 malfunction is distinguished from Type 3 because EIHC is normal, or only slightly high, and from Type 1 because there is no or only slight enrichment at moderate power.

There are seven vehicles of this type. Two vehicles, 42 (1990 Grand Am) and 71 (1992 Corolla), have burned-out catalysts. Five, 43 and 150 (both 1992 Dakotas), 77 (1992 Tercel), 254 (1992 Elantra), and 277 (92 Fox) are more complex examples of poor, highly-variable, catalyst performance; emissions characteristics for three of these vehicles are shown in Table 12. Vehicles 77 and 150 are similar in their relatively good fuel control and normal EIHC. Vehicle 43 and especially 254 and 277 have poor fuel control. Vehicle 277 could be classified as Type 2, with its considerable transient enrichment. Vehicle 254 could be classified as Type 3, being somewhat similar to 178; its EIHC is about twice normal.

Table 12. Emission Ratios for Type 4 (Vehicles 43, 77 & 150)

Variable	Range, Comment
EICO	up to 0.15, normal or slightly higher
EIHC	up to 0.025, normal or slightly higher
EINOx	< 0.05, normal
CPFCO	0.3 to 0.6, well above normal
CPFHC	» 0.2 or 0.3, above normal
CPFNOx	0.2 to 0.6, well above normal

The behavior of a vehicle with high emissions of all pollutants over a portion of the MEC (Figure 5a) is compared with that of a normal emitter (Figure 5b). Figure 5a illustrates strong if variable catalyst deterioration for vehicle 254, with CPFs of about 0.4 in moderate driving. This deterioration does not seem to be caused by excursions in phi, although we cannot be sure because the measurement of phi may not be accurate enough for this purpose. In contrast, Figure 5b shows that a normal emitter (vehicle 248) has CPFs of essentially zero in the same segment of the MEC (although CPFs do increase with excursions in phi).

The FTP bag 3 tailpipe emissions profile for all of these vehicles is shown in Table 13: in almost all cases all three pollutants are high, relative to clean car levels.

Table 13. FTP Bag 3 GPM Tailpipe Emissions for Type 4 Vehicles

Test Vehicle	CO	HC	NOx	profile
42 (car)	11.6	2.1	5.4	HHH
43 (truck)	10.4	0.7	2.5	HMH
71 (car)	9.2	1.6	1.9	HHH
77 (car)	7.1	1.0	1.7	HHH
150 (truck)	8.8	1.9	2.8	MHH
254 (car)	11.9	1.7	3.5	HHH
277 (car)	24.6	1.7	1.5	HHH

This type of high emitter may be associated with a burned-out catalyst, as observed in two of the vehicles here; but transiently bad catalyst performance is also observed. It is difficult to distinguish between two possible basic causes of the latter. The first involves greatly deteriorated performance of the catalyst, presumably due to severe operating conditions in the past. A second possible cause is poor closed-loop control of the fuel-air ratio, such that it doesn't conform to the needed pattern (illustrated in Figure 1), but at a level of failure too detailed to be observed directly here.

Summary

The CO/HC/NOx tailpipe emissions profiles for the 16 high-emitters measured in the NCHRP project and analyzed here, using the cutpoints of Table 4 to define the boundaries for High, Medium and Low, are shown in Table 14. We include MMH vehicles as both Type 1 and Type 4 high emitters, as discussed below.

Table 14. High-Emitter Types by FTP Bag 3 Profile

High-Emitter Type	CO/HC/NOx profile
1: lean	LLH, LMH, (MMH)
2: rich	HML, HMM
3: misfire	HHL, MHM, MHL, HHM
4: catalyst problem	HHH, MHH, (MMH)

An essential point is that these are general categories. Each "type" identified corresponds to more than one detailed behavior; for example, we observe both transient and chronic behavior for each type. And each type covers more than one disparate *physical* malfunction.

EMISSION PROFILES IN THE ARIZONA IM240 DATA

Because the high emitting vehicles recruited for testing under the NCHRP project are not representative of the in-use fleet, we analyze data from the Arizona I/M program to get a sense of the prevalence of each type of high emitter.

The IM240 test was recently introduced in several non-attainment areas, including the Phoenix area, as part of an enhanced inspection and maintenance (I/M) program. The test involves a 4-minute dynamometer cycle with speeds up to 57 mph, with an average speed of 30 mph. The IM240 power levels are similar to those in FTP bag 1 or 3, and involve the same maximum specific power, as shown in Table 1. To reduce costs and waiting, the 240-second test is terminated early by the Arizona contractor for vehicles with relatively low or high emissions. For short tests, we calculate an adjusted gpm; our adjustment is different than that used in Arizona [19].

DEVELOPMENT OF EMISSION PROFILES

Using the IM240 data, we create CO/HC/NOx profiles based on high, medium and low categories for each pollutant, as we did with FTP bag 3 measurements on the 16 NCHRP vehicles. The profiles again depend on choice of low-emitter and high-emitter cutpoints. (Because of differences between the two measurement programs, as discussed below, these IM240 cutpoints are not the same as those for the bag 3 measurements.) We consider several alternative sets of cutpoints; two of these sets, which differ in the definition of high-emitters, are shown in Tables 15 and 16.[1] Among MY1990-93 cars as measured in 1995, the cutpoints of Table 15 yield 10% high emitters (vehicles with at least one H); almost half of the non-high emitters are classified as LLL. The cutpoints of Table 16 yield 25% high emitters.

Table 15. High High-Cutpoints for Profiling the IM240 High Emitters

Range	CO (gpm)	HC (gpm)	NOx (gpm)
high H	>20	>1.2	>2.5
medium M	6 - 20	0.4 - 1.2	1.2 - 2.5
low L	<5	<0.5	<1.2

Table 16. Low High-Cutpoints for Profiling the IM240 High Emitters

Range	CO (gpm)	HC (gpm)	NOx (gpm)
high H	>15	>0.8	>2.0
medium M	6 - 15	0.4 - 0.8	1.2 - 2.0
low L	<5	<0.5	<1.2

Almost all of the Arizona IM240 high emitters occur in eight profiles, depending on the choice of cutpoints. The profile distributions found are shown in Table 17. With three pollutants and three emissions levels, H, M and L, there are nineteen possible profiles of high emitters (i.e. vehicles with at least one H). Just eight in Table 17 have an incidence of 5% or more; only 10% of the vehicles fall in the other eleven profiles. A characteristic of most of the missing profiles is that they do not obey a tight correlation between CO and HC (independent of the NOx level).

The distribution of a sample of vehicles among the high emitter profiles is shown in Figure 6. The vehicles all have at least one H, i.e. with one of the pollutants high. The dashed lines mark the boundaries of the emitter profiles, using the cutpoints in Table 15. The lower left quadrant of the figures represents the LLx emitter profile (low CO and HC, with unspecified NOx emissions), while the upper right quadrant contains cars in the HHx profile. The three level of NOx emissions are denoted in the figures using different symbols. One sees patterns: 1) There are no HLx and few LHx vehicles; i.e. HC and CO are strongly correlated. 2) High CO is correlated with low-to-moderate NOx. 3) There is a group of vehicles with high NOx and low-to-moderate CO and HC. These general tendencies are expected, but we are surprised by their pervasiveness in a very large sample. Part of the explanation is that high CO only occurs with enrichment, which enhances HC and suppresses engine-out NOx.

Care must be taken in interpreting the figure, since the restriction of at least one H strongly influences its appearance. Figure 7 is a similar scatterplot using the same cutpoints, but including vehicles with two medium-level pollutants, in order to clarify the structure near the medium-to-high transition in HC for medium CO. The distribution is smooth across this boundary. One sees, for example, that there are many MML vehicles, with medium CO, but on the high side, which probably have similar malfunctions to those classified as HML, i.e. with high CO.

Table 17. Distribution of High Emitters by Profile: Arizona IM240, MY1990-1993 Cars[a]

Profile: CO/HC/NOx	Percent high emitters: high cutpoints[b]	low cutpoints[c]
HHH	1	3
HHM	5	5
HMH	0	0
MHH	11	18
HMM	2	1
MHM	17	12
MMH	20	10
HHL	10	10
HML	11	5
HLM	0	0
MHL	6	8
MLH	2	4
LHM	1	2
LMH	4	4
HLH	0	0
LHH	0	2
HLL	0	2
LHL	0	1
LLH	7	13

a) since we base the emission profile on our adjusted gpm results from the IM240 data, some cars classified as high emitters in this analysis actually were passed by the AZ I/M contractor (were passed in Phase 2 of test).
b) See Table 15. c) See Table 16.

FREQUENCY OF OCCURRENCE OF TYPES OF HIGH EMITTERS

All but three of the eight important IM240 profiles (Table 17) are included in the list of profiles identified among the NCHRP/Riverside high emitters (Table 7); the three are MMH, LMH and MHL. The differences between the two sets of percentages in Table 17 show where there are sensitivities to the high cutpoints used.

1 . The high cutpoints shown in Table 15 are the cutpoints currently in use in the Arizona I/M program for MY1991 and newer cars. The high cutpoints in Table 16 are the final cutpoints originally proposed for the Arizona program (and not adopted due to the finding of inconsistent vehicle preconditioning [13]).

High emitters from the NCHRP project (FTP bag 3) are plotted in Figure 8 for comparison with the sample of the Arizona IM240 high emitters in Figures 6 and 7. Figure 8 has the same axis scales as Figures 6 and 7, but the dashed lines reflect the lower cutpoints used for the FTP tests.

In Figure 9 we present rough boundaries for the IM240 profiles for the four types of high emitter identified among the NCHRP vehicles. As seen, we assign about one-third of IM240 category MMH to Type 4 and two-thirds to Type 1, all of LMH to Type 1, and all of MHL to Type 3. The resulting frequencies as percentages of all high emitters are shown in Table 18.

Table 18. Distribution of IM240 Profiles of MY90-93 Cars, Based on Cutpoints of Table 15

High Emitter Type	Profile	Percent of High Emitters	Percent of All Cars
1: Runs Lean	LLH, LMH, (MMH)	24	2.4
2: Runs Rich	HML, HMM	13	1.3
3: Misfire	HHL, MHM, MHL, HHM	38	3.8
4: Bad Catalyst	HHH, MHH, (MMH)	19	1.9
Other high emitters		5	0.5

CAVEAT

There are several important differences between IM240 bag emissions as measured and those of FTP bag 3 analyzed above:

- The sample of vehicles is quite different. IM240 test results of over 135,000 MY90-93 passenger cars were analyzed; these vehicles represent roughly half of the registered vehicles in the Phoenix area (the program is a biennial program, where testing is required every two years and upon vehicle sale). These data are much more representative of the in-use fleet than the 300 vehicles tested under the NCHRP program. In addition, the Arizona data are dominated by 49-state vehicles with somewhat different emissions controls than for California vehicles. Moreover, the measurements in Arizona used here were made in 1995, while those at UC Riverside were made in 1996-97. In addition, the IM240 sample used consists of cars only, while the NCHRP data contains both cars and light trucks.
- The conditioning of the vehicles (i.e. the block and catalyst temperatures prior to testinig) is somewhat different. This is probably not a big effect for high emitters. As an extreme comparison, when one compares the NCHRP FTP bag 2 and bag 3 data one finds that bag 2 HC and CO emissions are only moderately lower, in spite of the full warm-up and lower power requirements of bag 2.
- Most important, we are comparing carefully controlled FTP measurements carried out on 300 vehicles in a laboratory setting with relatively inexpensive measurements on over one hundred thousand vehicles. The equipment and procedures are different; and the CE-CERT group at Riverside has found that it is not a routine matter, even in their laboratory setting, to obtain accurate results. We find that the Arizona IM240 measurements tend to exaggerate the emissions of low- and medium-emitting vehicles, a subject we will explore in a different report. (This does not mean that the Arizona measurements fail to satisfy their purpose, the identification of high emitters.)
- Another problem with the IM240 analysis is that about half of the IM240 tests analyzed were ended after 31 seconds of driving, because the cars met low "fast pass" emission cutpoints. And most of those tested more than 31 seconds were also given a shortened test. Only about 2% of the tested cars were given the full IM240 test; most of these cars were randomly recruited to receive the full test. Although we make adjustments to make the shortened test emission results roughly comparable to those of a full IM240 test, these adjustments are rather simplistic and may affect our results.

All of these differences between the FTP and IM240 testing may affect the accuracy of mapping FTP high emitter types to IM240 emission profiles.

DISCUSSION

Generally speaking, the four types of high emitters identified from the emission ratios are roughly equally represented in the Arizona I/M fleet. Type 1 (runs lean) occurs in 24% of vehicles while Type 2 (runs rich) occurs in only 14%. It is possible that there has been a shift in the distribution of high emitters from high CO to high NOx emitters, as we have moved from carbureted to sophisticated computer-controlled fuel-injected vehicles. Also, earlier I/M programs using idle emissions tests virtually ignored NOx emissions, so high emitters may have been previously repaired to reduce CO and HC at the expense of NOx emissions.

For many people, the study of emissions-control malfunction concerns component malfunction. While our study does not directly address individual components, we do get some information on what components may affect the different types of high emitters. As just mentioned, we find that relatively small fuel control deviations from stoichiometry characterize about 40% of the high emitters. Another group (33%) can be roughly characterized as cylinder misfire (Type 3). Catalyst malfunction in the absence of one of the other malfunctions (Type 4) has a relatively low probability at 19%. However, catalyst malfunction is an important but subsidiary problem in many Type 2 and 3 vehicles. So the statement that replacing the catalyst will improve the emissions performance in one-half or more of vehicles is in agreement with our data. But the improvement might be temporary in many vehicles because uncorrected conditions of frequent enrichment or misfire might cause swift catalyst degradation.

In the NCHRP sample, we did not find excessive lean operation to be associated with catalyst degradation. We have not gone further in attempting to pinpoint component failures from the NCHRP data. The data are rich; we hope that others will study it to discover more.

LIMITATIONS

There are several analytical and measurement limitations to this study. Most have been mentioned, but they are worth a reminder: a) Accurate measurement of fuel-air ratio is difficult, so much of what we conclude about this critical aspect of emissions control is inferred. b) The sample of NCHRP vehicles is small, and has been further sliced into many categories. To the extent study results are as important as we think they are, this study should be followed up by one with substantially more tests of modern high-emitters. c) Most of the measurements involved MY1990-93 vehicles, which we have treated as a group. We have not examined changes in vehicular emissions control technologies during the 1990s. d) The use of profiles involves cutpoints, with the attendant sensitivity to choice of cutpoints. We have examined a few sets of cutpoints for the IM240 data and find that the general results hold for these cutpoints. e) Verification of the accuracy of the IM240 measurements at high gpm levels needs to be improved.

APPLICATIONS

The application that led to this work as part of the NCHRP project is the inclusion of high emitters in modal emissions modeling, i.e. inclusion of the dependence on driving pattern of emissions from malfunctioning vehicles. What we have been able to do is a first step. The sample of NCHRP high emitters from MY90 and later is inadequate to accurately determine modeling parameters for the four types each with chronic and transient subclasses. We can nevertheless use a weighted mix of the measured vehicles to create a detailed simulation model of emissions as they depend on operating variables such as speed, acceleration and grade. As an example of what might be found, we note that high emitters of Types 1, 2 and 3 may be less sensitive to power than to transients, while for Type 4 power is the key operating variable.

While a first step, such modeling of high emitters would constitute a major improvement in modal modeling; and it should also contribute to emissions inventory modeling. An issue of interest not yet been studied, but accessible in the NCHRP data, is emissions from modern high emitters at high power levels (beyond the FTP range).

Another application is to help achieve more-durable emissions control through the categorization. The three-pollutant profiles obtained in high-statistics and low-bias recruitment measurement surveys may enable one to focus on important high-emitter problems among recent vehicles. For example, through this research we have begun to be able to a) accurately assess the role of throttle fluctuation and driving with frequent speed adjustment, and b) throw light on catalyst degradation as a result of failure of other controls in contrast to severe driving.

In this paper our focus is categorization of high emitters. We do not address the issue of the total contribution of high-emitting modern vehicles to the emissions inventory. This result depends on assumed cutpoints. A full and fair evaluation of the role of modern high emitters in the emissions inventory is critically important, and requires a different study than the categorization analysis carried out here.

FINAL COMMENT

We believe that systematic measurement surveys with high-statistics and low-bias recruitment could be extremely useful for programs to assess in-use durability of emissions controls in modern vehicles. Such surveys could be based on IM240, remote sensing, on-board diagnostics or some other technique. Until now, in-use testing programs by regulatory agencies and the manufacturers have been severely weakened by possible biases in recruiting high emitters and by poor statistics. As a result of these problems, the nation does not have convincing evidence one way or the other on the importance of high emitters among modern vehicles. We believe that careful analysis of I/M data collected by states can shed light on the real-world emissions of modern vehicles.

CONCLUSION

In this study we examine second-by-second pollution outputs, including engine-out emissions, of vehicles which are high-emitters in low-to moderate-power driving (within the FTP range). We use these detailed emissions data to infer possible causes of emission control system malfunction.

We observe four different patterns, or types, of emissions control malfunction: 1) fuel-air ratio excessively lean, 2) fuel-air ratio excessively rich, 3) partial combustion such as misfire, and 4) severe deterioration in catalyst performance in vehicles where malfunctions of Types 1, 2 or 3 are not predominant. For many vehicles, more than one malfunction is observed; we characterize the malfunction by the one which is the most important to the high emissions. In addition, for all four types of high emitter two further categories are observed: transient and chronic. Transient high emitters are extremely sensitive to vehicle speed variations, or throttle fluctuations; their emissions control performance may be good in steady low-power driving. Chronic high emitters have roughly steady patterns of emissions control failure.

We then relate the types of high emitters, as defined by the analysis of emissions ratios, with 3-pollutant profiles of tailpipe emissions (expressed as high, medium, or low tailpipe emissions of CO, HC and NOx, or, for example, HLM). The correspondences allow us to relate the detailed analysis of a small number of high emitters to bag data for a large number of vehicles tested in the IM240 program in the Phoenix Arizona area, to determine the distribution of high emitter types within the on-road fleet. The emission cutpoints are chosen so that the resulting emission profiles are consistent with the emissions ratios; however, we cannot definitively demonstrate the validity of using tailpipe emissions alone to characterize high emitter types.

We find that CO and HC emissions are correlated; if one is high the other is not low. And CO and NOx are negatively correlated; if one is high, the other is not high. All four types of high emitters--improper fuel control (lean or rich operation), misfire, and catalyst deterioration--are observed in the NCHRP testing, and are roughly equally represented in the Arizona I/M fleet.

ACKNOWLEDGMENTS

The authors would like to thank their colleagues in the NCHRP project at CE-CERT, UC Riverside, including Feng An, Matthew Barth, Jim Morgan, Joseph Norbeck, George Scora, Timothy Truex, and Theodore Younglove, as well as professor Robert Sawyer of UC Berkeley, John DeCicco of the American Council for an Energy-Efficient Economy, Harold Haskew of General Motors, and John German of Honda. The research reported here was partially supported by the NCHRP, Project 25-11.

REFERENCES

1. Donald H. Stedman, "Automobile Carbon Monoxide Emission," *Envir. Science Technol.* 1989, **23**, 147.
2. D.R. Lawson, P. Groblicki, D.H. Stedman, G.A. Bishop, and P.L. Guenther, "Emissions of In-Use Motor Vehicles in Los Angeles: A Pilot Study of Remote Sensing and the Inspection and Maintenance Program," *J. Air & Waste Manage. Assoc.*, **40**, 1096-1105, 1990.
3. Robert D. Stephens, "Remote Sensing Data and a Potential Model of Vehicle Exhaust Emissions," *J. Air & Waste Manage. Assoc.*, **44**, pp. 1284-1292, 1994.
4. California Air Resources Board, *On-Road Remote Sensing of CO and HC emissions in California*, prepared by Donald Stedman, Gary Bishop, Stuart Beaton, James Peterson, Paul Guenther, Ian McVey, and Yi Zhang, CARB Research Division, Contract No. A032-093, February 1994.
5. California Air Resources Board, *Comparison of the IM240 and ASM Tests in CARB's I&M Pilot Program*, CARB Mobile Source Division, June 1996.
6. Ken McAlinden, "Michigan Roadside Study: Analysis of Repairs on High Emission Vehicles," from *Proceedings of the 4th CRC On-Road Vehicle Emissions Workshop*, Coordinating Research Council, Atlanta GA, 1994.
7. Ahmed Soliman and Georgio Rizzoni, "The Effects of Various Engine Control System Malfunctions on Exhaust Emissions Levels During the EPA I/M 240 Cycle," Society of Automotive Engineers, Warrendale PA, 940448, 1994.
8. Philip L. Heirigs and Thomas C. Austin, "Causes of Failure in High Emitting Cars," Society of Automotive Engineers, Warrendale PA, 961280, 1996.
9. American Petroleum Institute, *Analysis of Causes of Failure in High Emitting Cars*, API Publication Number 4637, February 1996.
10. Feng An, Matthew Barth, Joseph Norbeck, and Marc Ross, "The Development of a Comprehensive Modal Emissions Model: Operating Under Hot Stabilized Conditions," Transportation Research Record no. 1587, pp 52-62, 1997.
11. Matthew Barth, Theodore Younglove, Tom Wenzel, George Scora, Feng An, Marc Ross and Joseph Norbeck, "Analysis of Modal Emissions from Diverse In-Use Vehicle Fleet," Transportation Research Record no. 1587, pp 73-84, 1997.
12. Donald H. Stedman, Gary A. Bishop, Philip Aldrete, and Robert S.Slott, "On-Road Evaluation of an Automobile Emission Test Program," *Envir. Science Technol.* 1997, **31**, 927-931.
13. Philip L. Heirigs and Jay Gordon, "Preconditioning Effects on I/M Test Results Using IM240 and ASM Procedures," Society of Automotive Engineers, Warrendale PA, 962091, 1996.
14. Robert W. Goodwin and Marc Ross, "Off-Cycle Exhaust Emissions from Modern Passentger Cars with Properly-Functioning Emissions Controls," Society of Automotive Engineers, Warrendale PA, 960064, 1996.
15. Robert W. Goodwin, *A Model of Automobile Emissions During High-Power Driving Episodes and Related Issues*, Ph. D. thesis, University of Michigan 1997.
16. Feng An, Matthew Barth, George Scora, and Marc Ross, "Modeling Enleanment Emissions for Light-Duty Vehicles," to be presented at the 77th Annual TRB meeting, Washington DC, January 1998.
17. Harold M. Haskew, Kevin Cullen, Thomas F. Liberty, and William Langhorst, "The Execution of a Cooperative Industry/Government Exhaust Emission Test Program," Society of Automotive Engineers, Warrendale PA, 94C016, 1994. (CD-ROM of data can be obtained from Jim Markey at the USEPA National Vehicle and Fuel Emissions Laboratory, 2565 Plymouth Road, Ann Arbor, MI 48105).
18. Harold M. Haskew, Doug Berens, and Roger Orteca, "FTP Emissions from 1991-1993 MY In-Use Vehicles," from *Proceedings of the 7th CRC On-Road Vehicle Emissions Workshop*, Coordinating Research Council, Atlanta GA, 1997. (Data can be obtained from the USEPA National Vehicle and Fuel Emissions Laboratory, 2565 Plymouth Road, Ann Arbor, MI 48105).
19. Tom Wenzel, "Analysis of Emissions Deterioration of In-Use Vehicles, Using Arizona IM240 Data," presented at the Society of Automotive Engineers Government/Industry Meeting, Washington DC, May 1997.

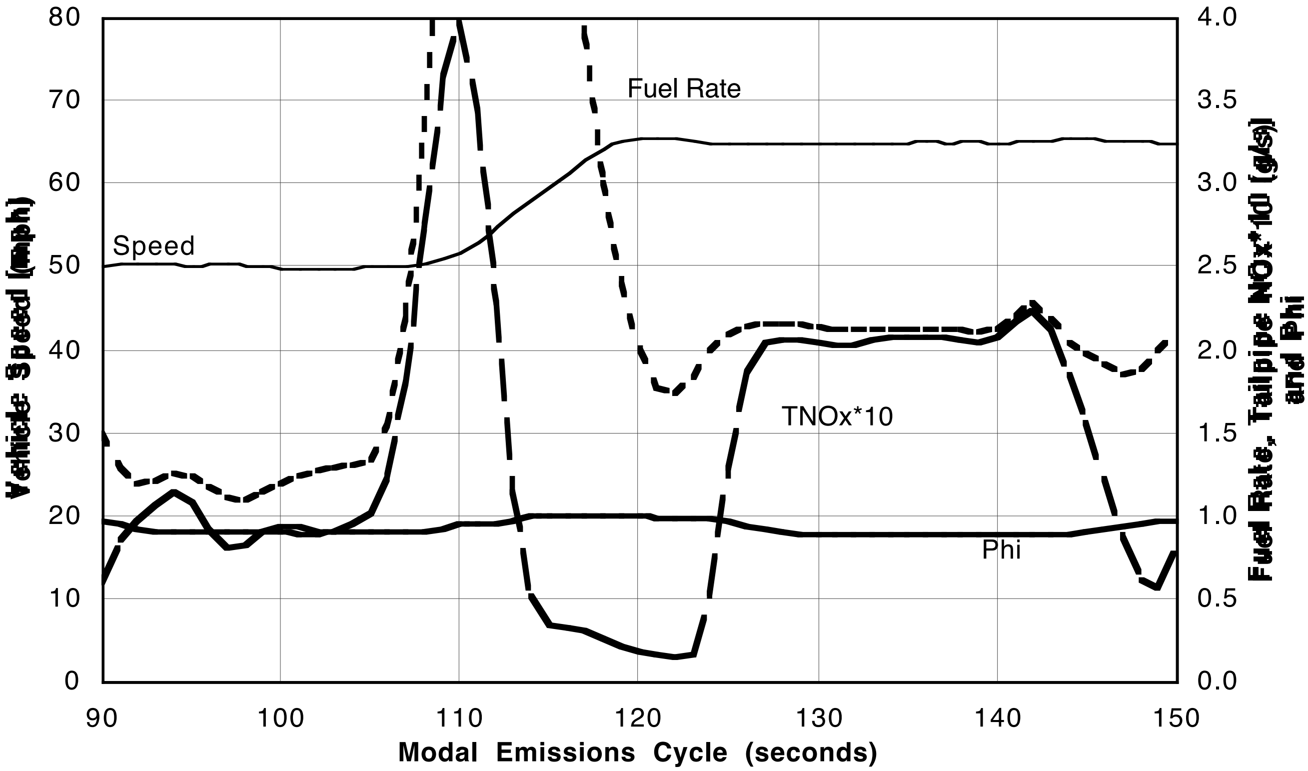


Figure 2a. Vehicle 202 (High NOx Emitter): Fuel Rate, Tailpipe NOx, and Phi

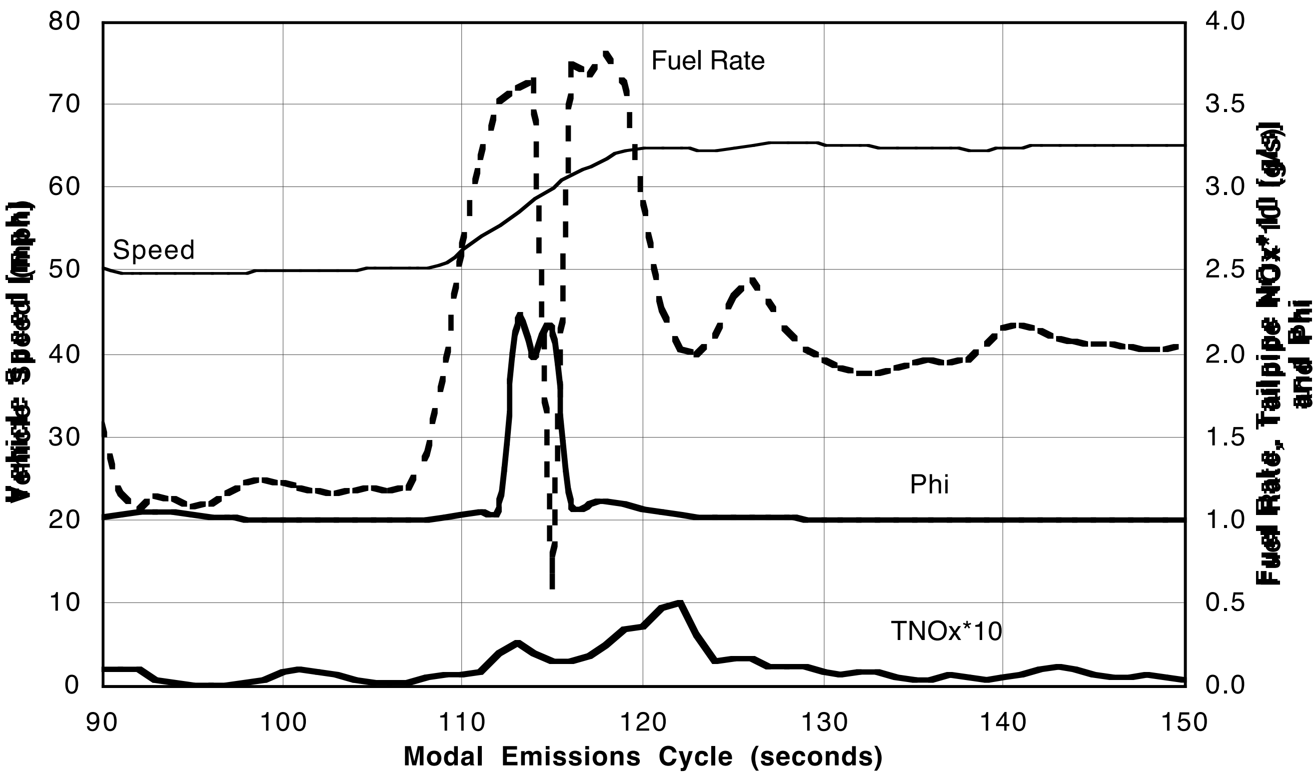


Figure 2a. Vehicle 136 (Normal NOx Emitter): Fuel Rate, Tailpipe NOx, and Phi

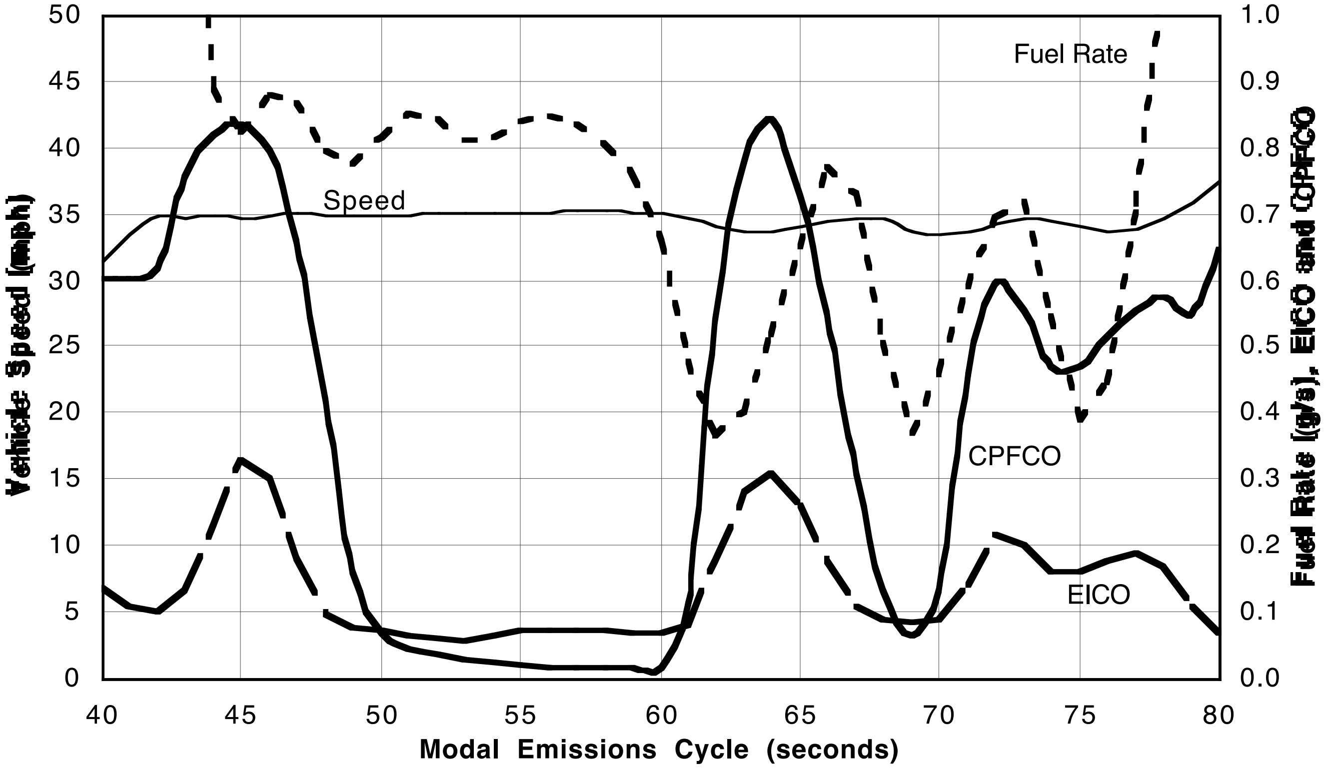


Figure 3a. Vehicle 136 (High CO Emitter): Fuel Rate, Engine Out CO and CO Catalyst Pass Fraction

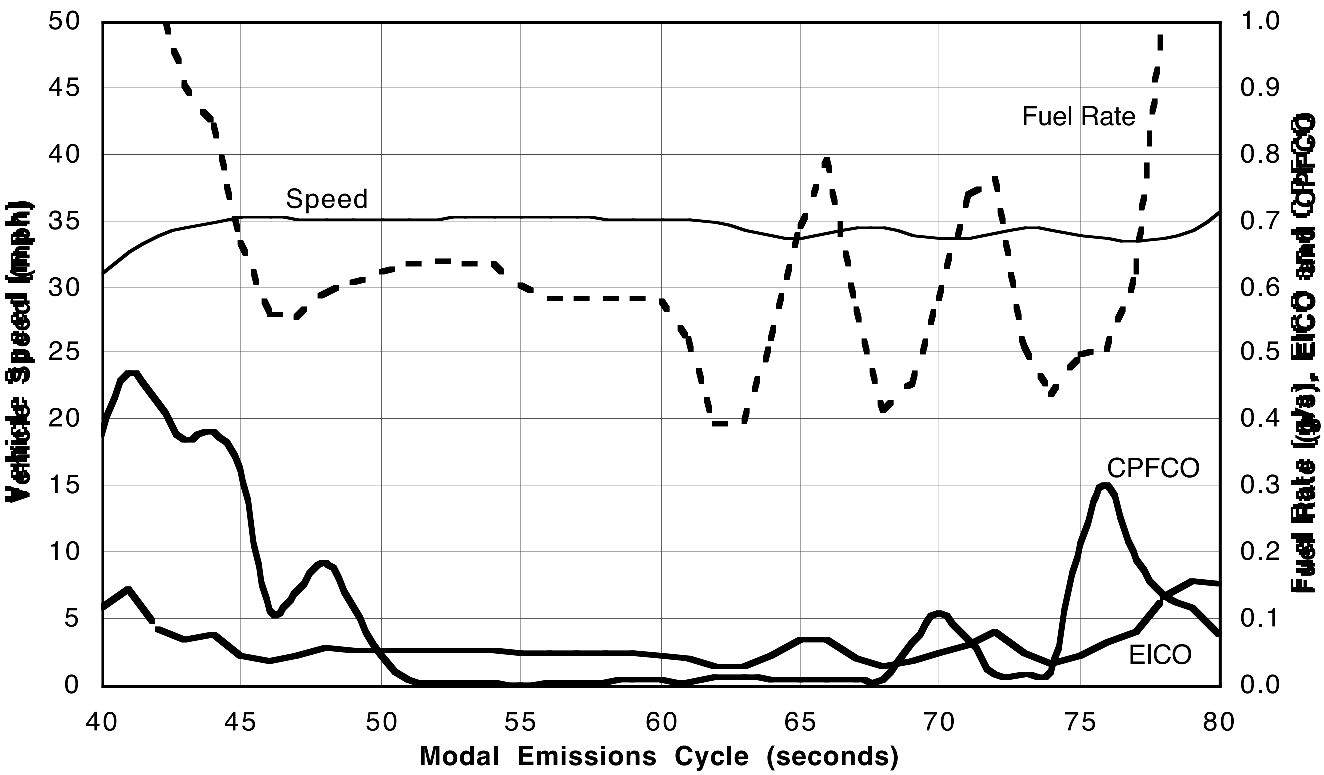


Figure 3b. Vehicle 103 (Normal CO Emitter): Fuel Rate, Engine Out CO and CO Catalyst Pass Fraction

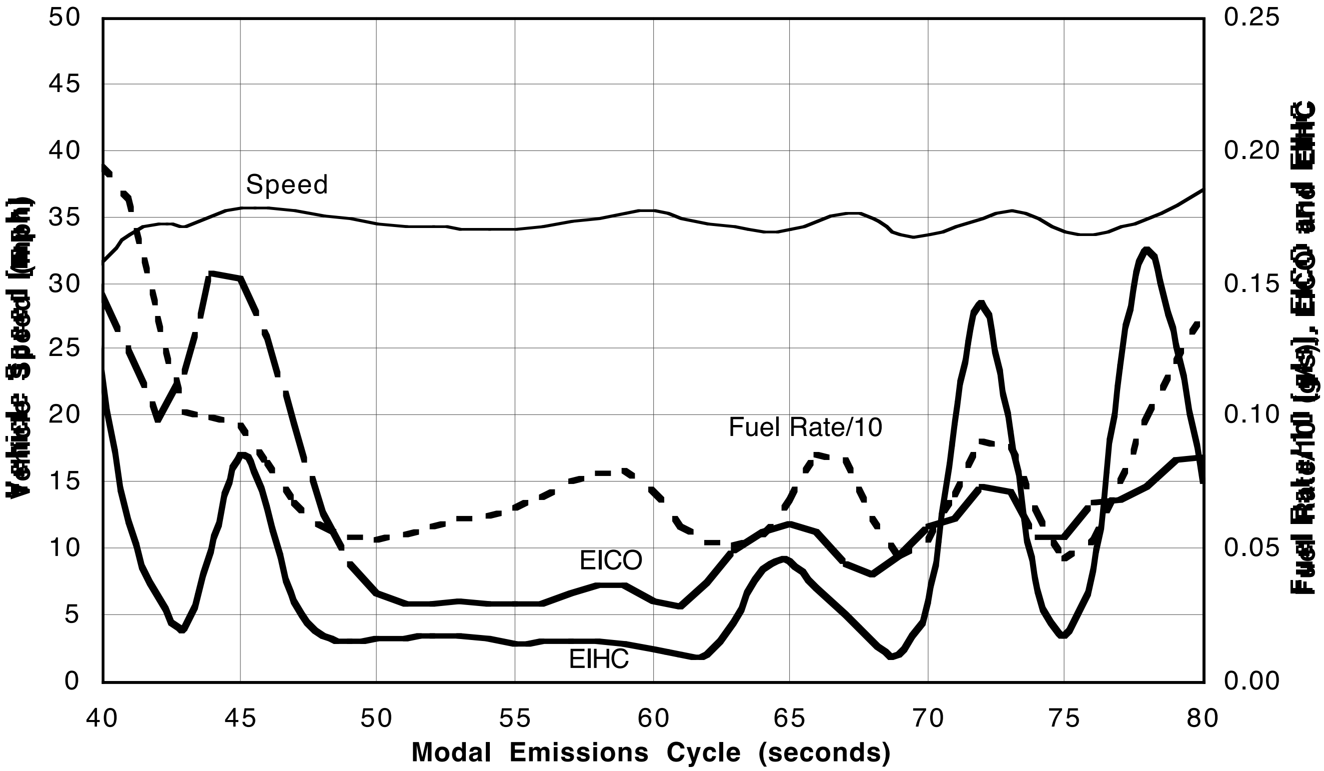


Figure 4a. Vehicle 178 (High HC Emitter): Fuel Rate, Engine Out CO and HC

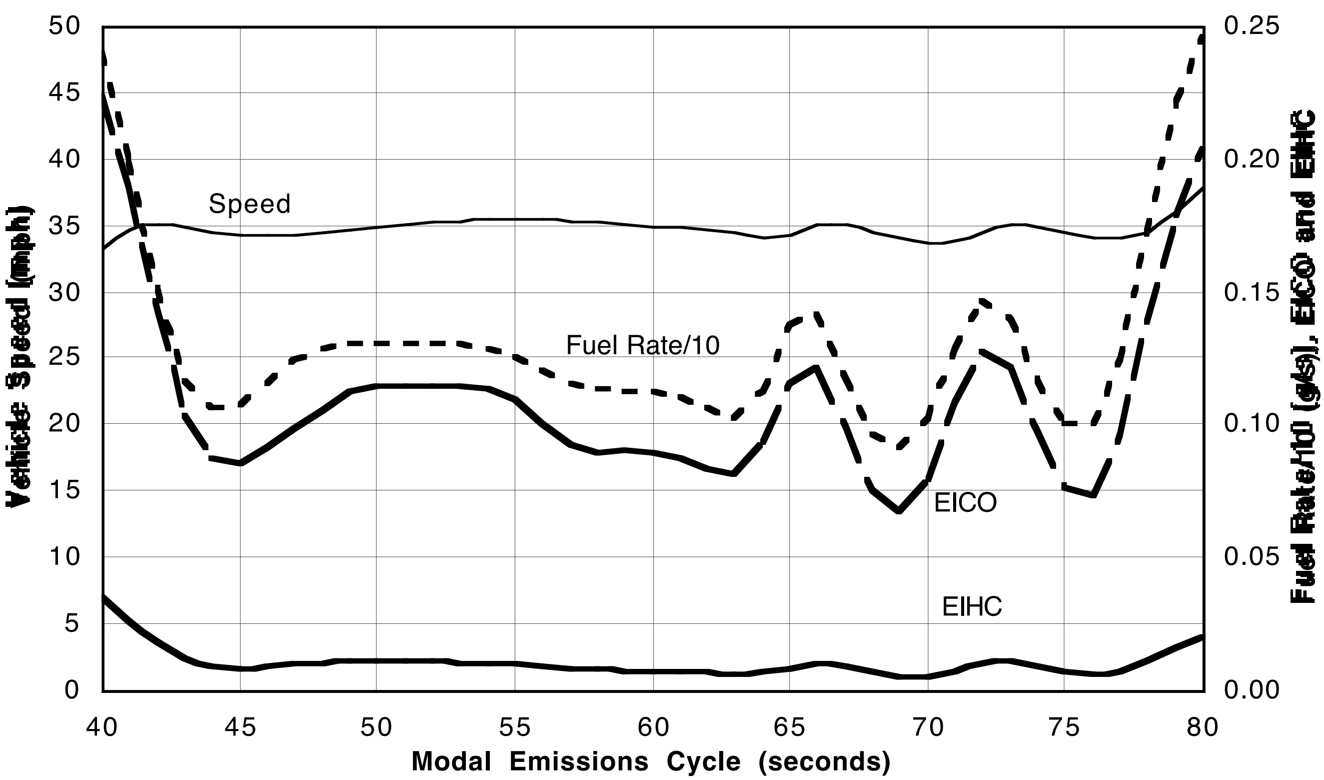


Figure 4b. Vehicle 295 (Normal HC Emitter): Fuel Rate, Engine Out CO and HC

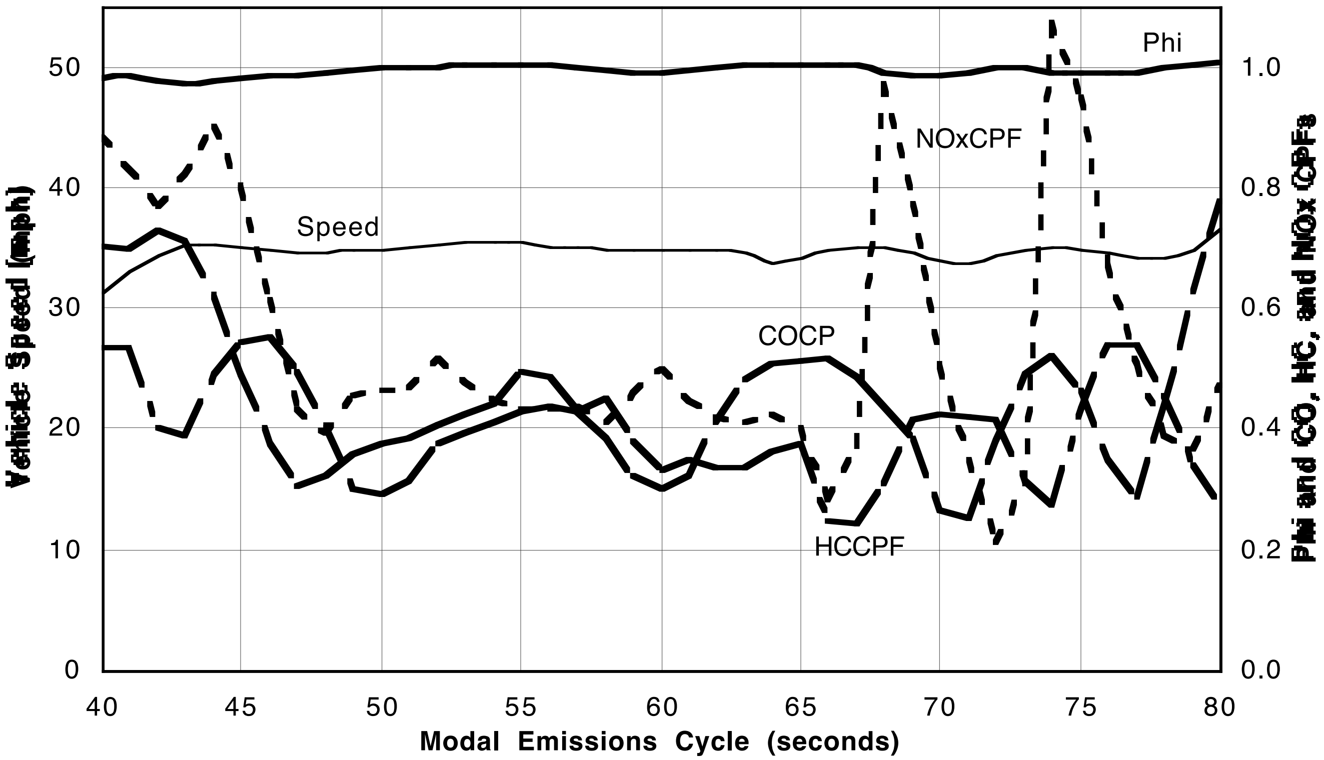


Figure 5a. Vehicle 254 (High CO, HC, and NOx Emitter): Phi and CO, HC, and NOx CPFs

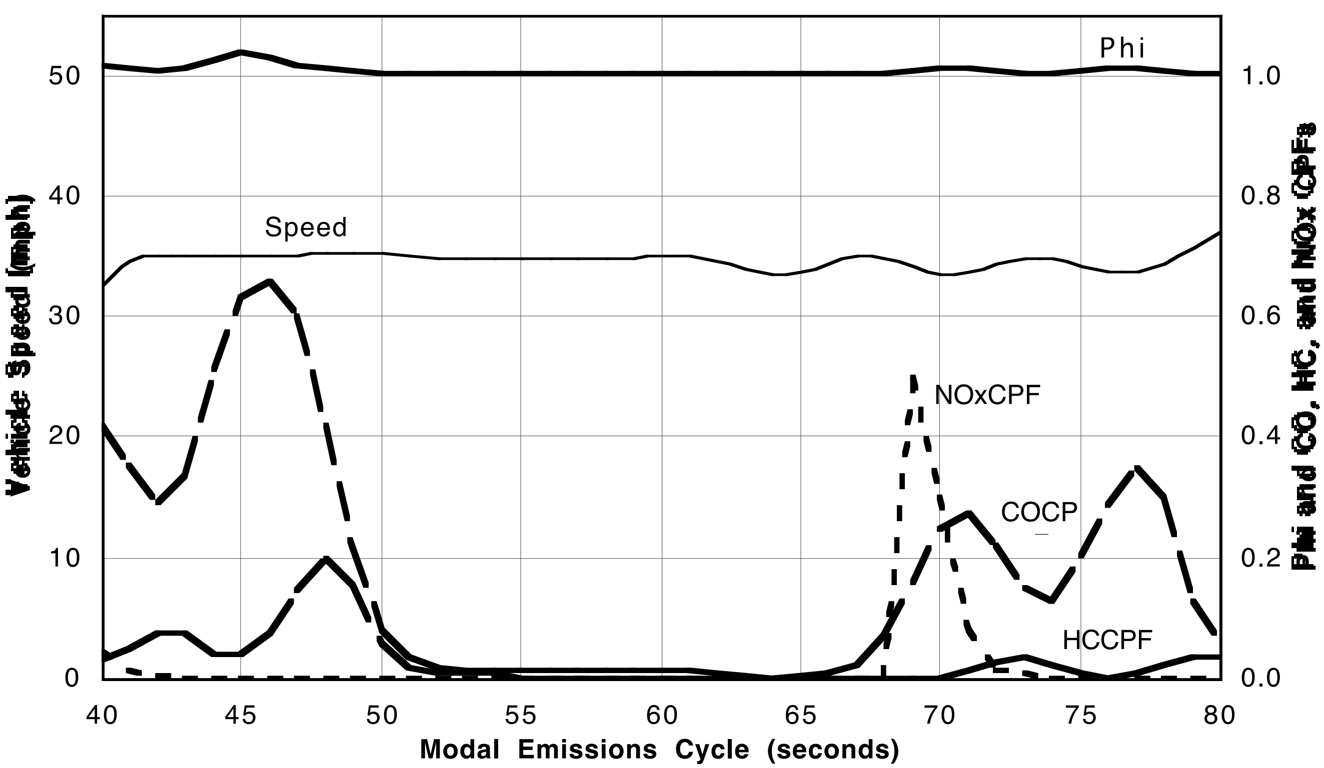


Figure 5b. Vehicle 248 (Normal CO, HC, and NOx Emitter): Phi and CO, HC, and NOx CPFs

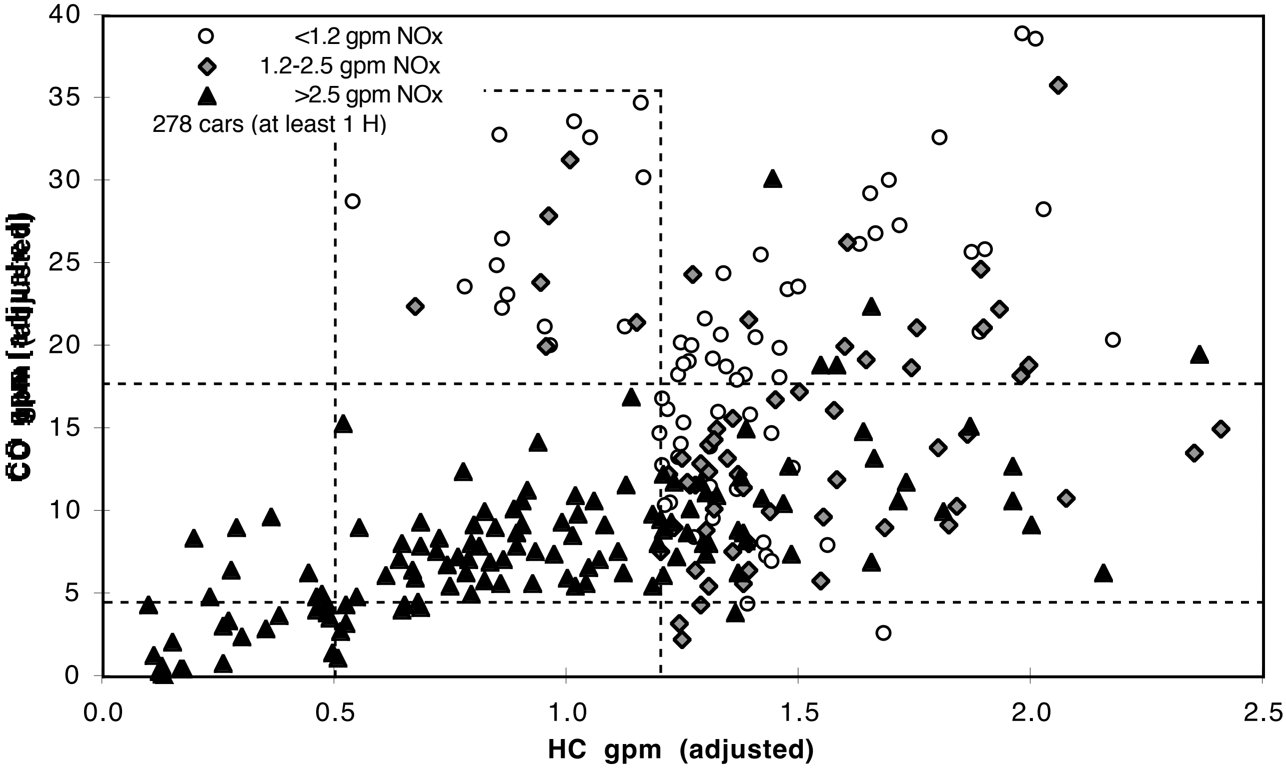


Figure 6. Distribution of High Emitters by Emission Profile (CO/HC/NOx), 278 Cars with at Least 1 H (MY90-93 Cars, 1995 AZ IM240)

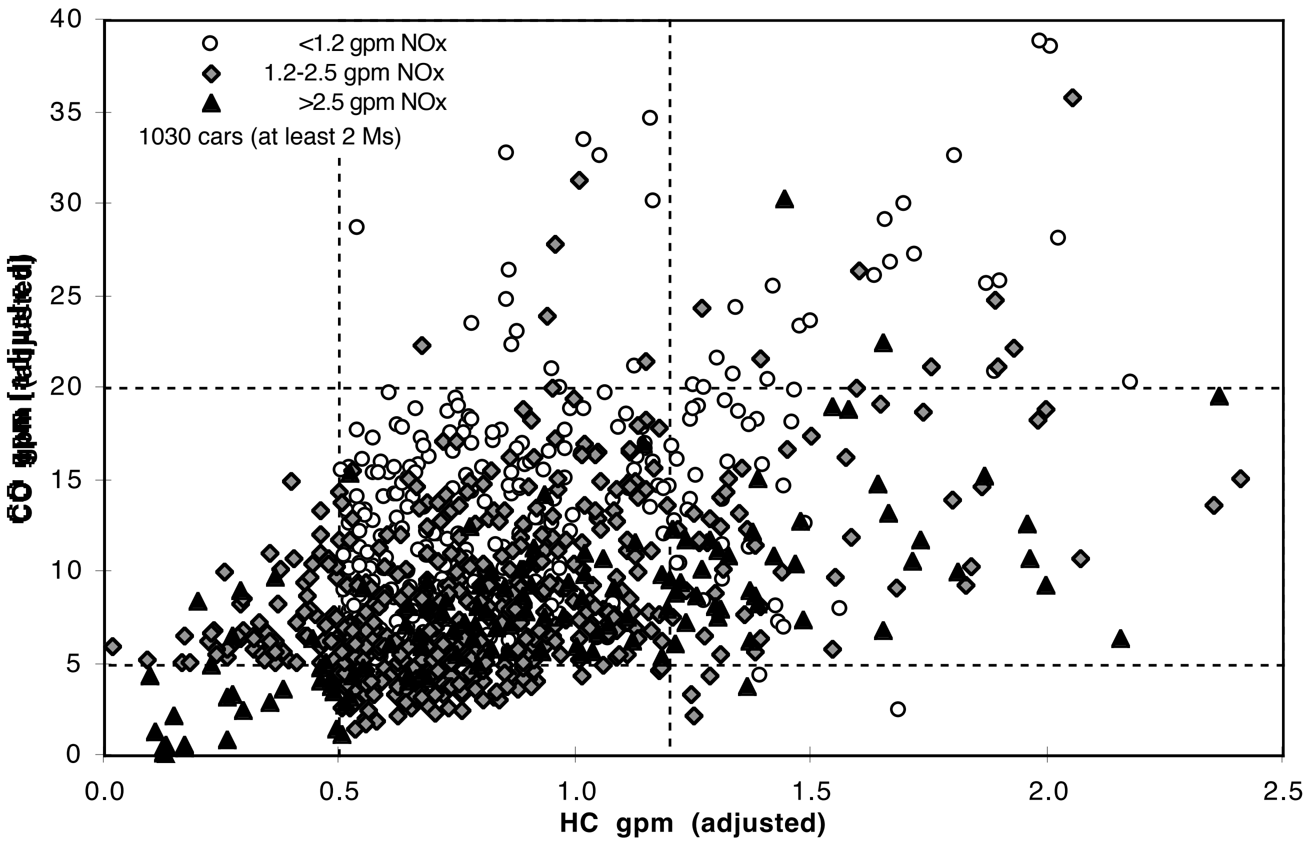


Figure 7. Distribution by Emission Profile (CO/HC/NOx), 1030 Cars with at Least 2 Ms (MY90-93 Cars, 1995 AZ IM240)

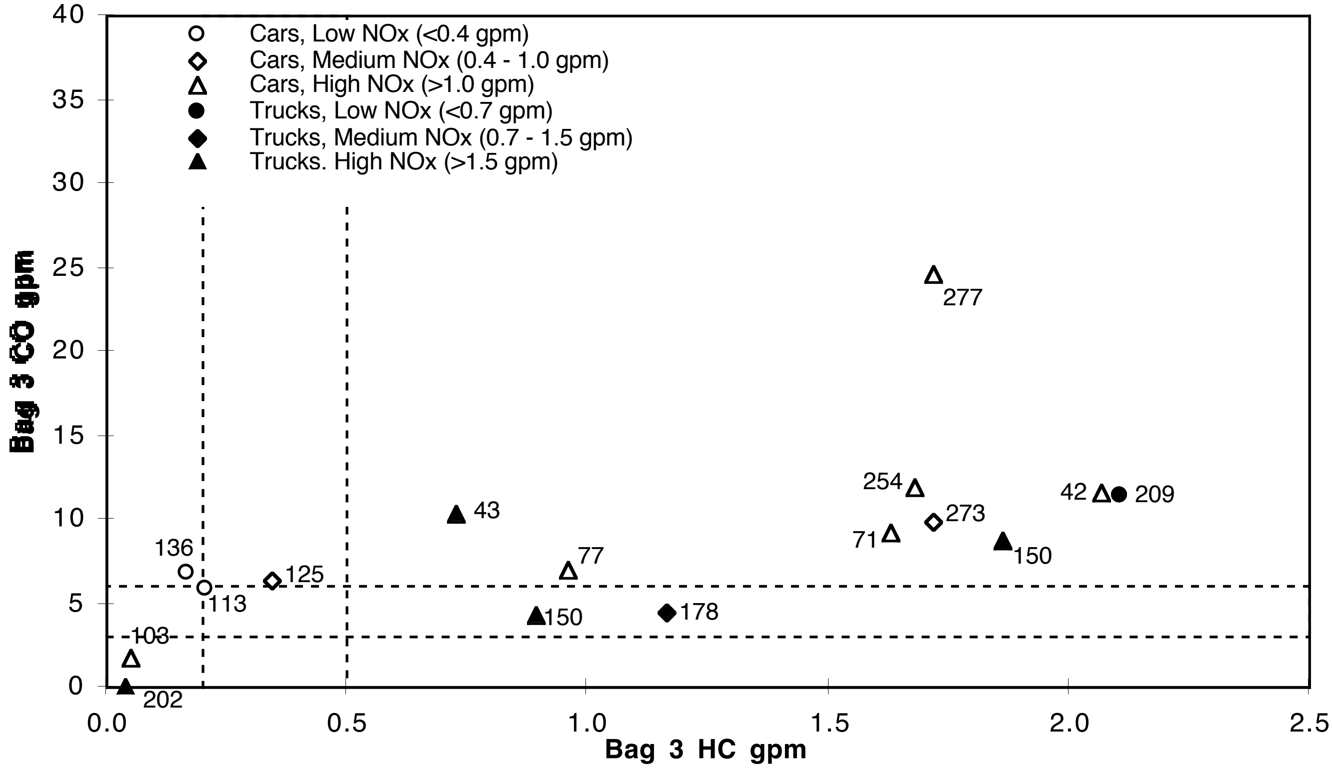


Figure 8. Distribution by High Emitter Type, MY90-96 Vehicles, 1996-97 NCHRP FTP

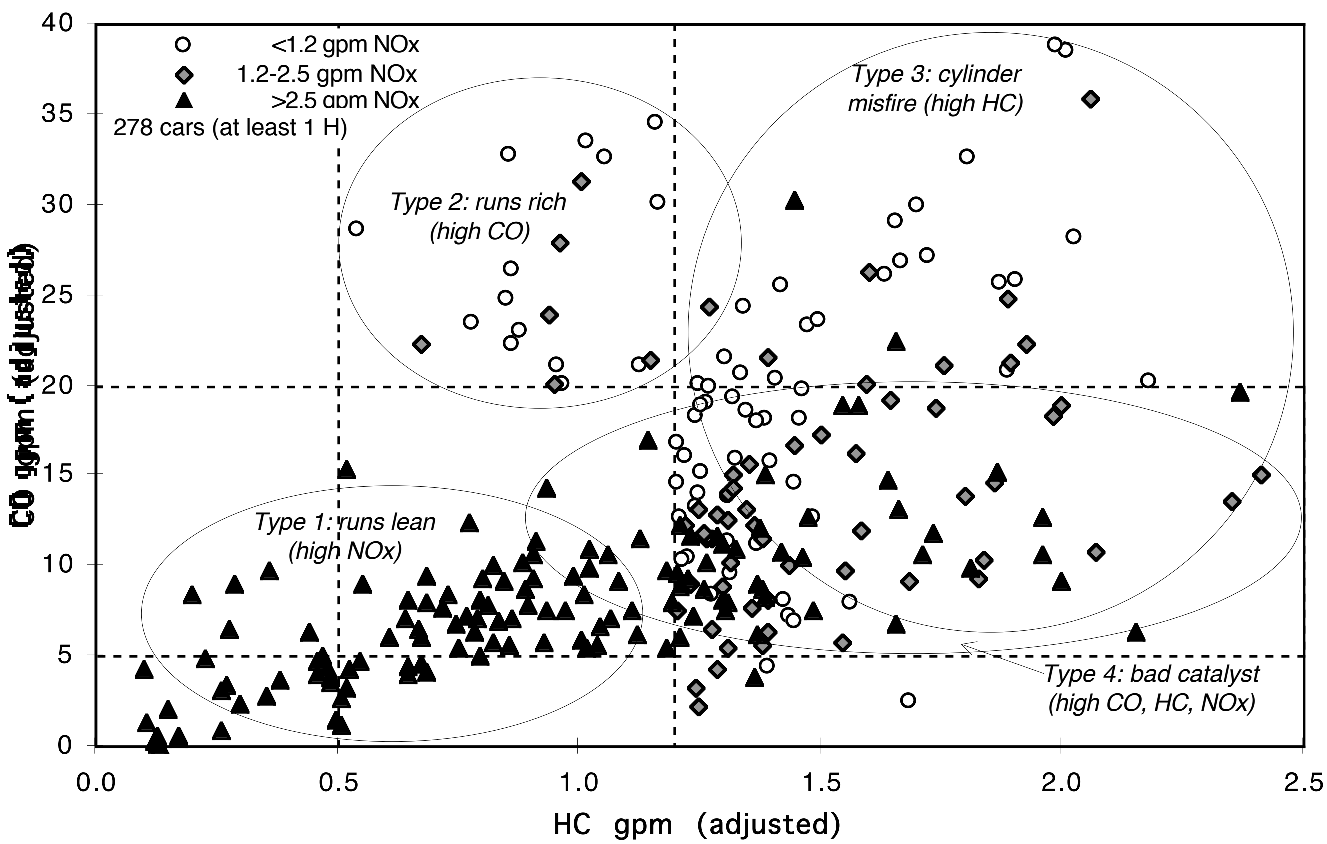


Figure 9. Distribution by Emission Profile and Type, MY90-93 Cars, 1995 AZ IM240

APPENDIX T

10TH CRC ON-ROAD VEHICLE EMISSIONS WORKSHOP
San Diego, California
March 27-29, 2000

COMPARISON OF THE INCIDENCE OF HIGH EMITTERS IN IM240 AND REMOTE SENSING MEASUREMENTS

Marc Ross[a] and Tom Wenzel[b]

a) University of Michigan, Physics Dept., Ann Arbor MI 48109-1120
b) Lawrence Berkeley National Laboratory, 90-4000, Berkeley CA 94720

ABSTRACT

Very large IM240 and/or remote sensing surveys of vehicle emissions are conducted in several states. These data are primarily obtained to identify individual vehicles that fail emissions criteria, with the owner asked to engage in a follow-up process of repair and retesting. However, these data may also enable an accurate assessment of the contribution of high emitters to the automotive emissions inventory. Since very large numbers of vehicles are measured, and the "recruitment" may be relatively unbiased, important statistical information on the emissions performance of selected vehicle cohorts may be obtained. This presentation focuses on vehicles driven over 100,000 miles to evaluate the major trends in the contributions of high emitters, and it briefly examines some strengths and weaknesses of the two measurement techniques for this purpose.

Analysis of large IM240 datasets from Arizona, Colorado and Wisconsin, show some important features: The contribution of high emitters rises with odometer reading, declines with model year at fixed odometer reading, and is independent of vehicle age for fixed odometer reading and model year. However, there are major questions about recruitment bias.

To check the results from the IM240 surveys, we examine remote sensing measurements in Arizona on a set of vehicles also measured in the IM240 (thus obtaining approximate mileage). The RSD emissions distributions agree rather well with those from the IM240. In particular, the declines in the high-emitter contribution with model year at fixed odometer reading are similar; and the comparison in terms of pollution indices, shows good agreement in spite of the differences that must characterize the two surveys, including differences in power output.

COMPARISON OF THE INCIDENCE OF HIGH EMITTERS IN IM240 AND REMOTE SENSING MEASUREMENTS

This presentation is organized around three main results:

1) IM240 surveys show that the incidence and contribution of high emitters declined rapidly with model year for late-80s through early-90s for cars in fixed high-mileage groups. In other words, improvements in technology have greatly reduced the incidence of malfunctioning emissions controls.
2) However, the use of IM240 surveys for statistical purposes like this is questionable because of measurement and recruitment problems.
3) Initial analysis of remote sensing surveys confirms the sharp decline of high emitters with model year, although the decline is not as great in the RSD for model years 1993-95 as in the IM250.

High Emitters and "Vehicle Probability Distributions". The main analytical tool used here to analyze the incidence and contribution of high emitters is the **vehicle probability distribution, or** VPD. VPDs enable one to visualize the role of high emitters among all vehicles in a cohort, as explained by Figures 1 and 2.

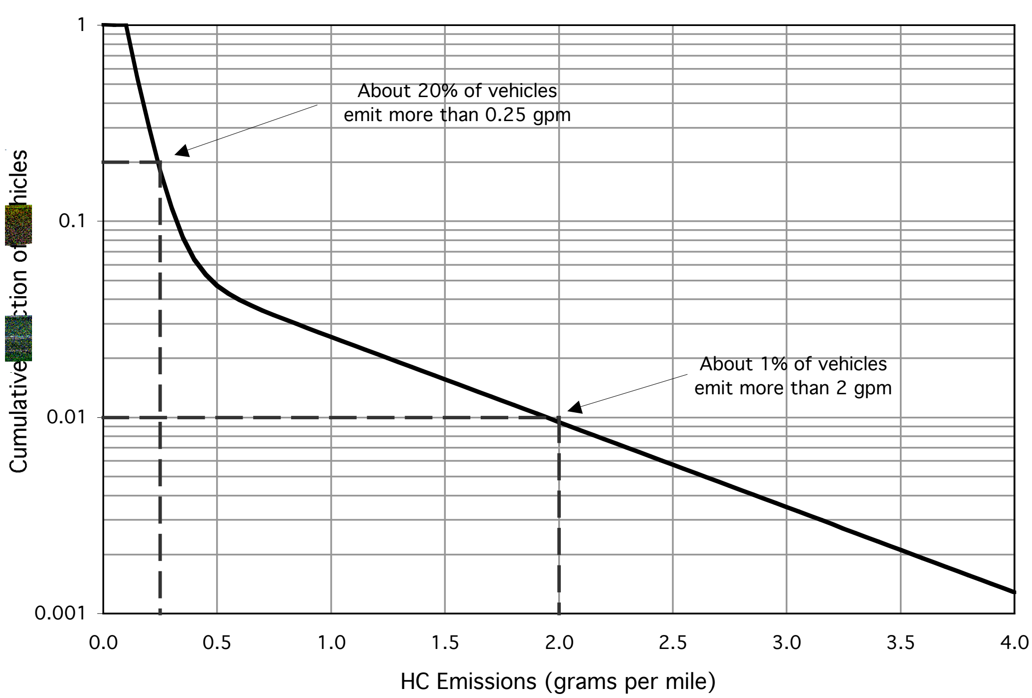


Fig. 1. Definition of the Vehicle Probability Distribution (fraction of vehicles with emissions > x). The sketch roughly shows HC IM240 emissions by MY1991 cars driven 60-100k miles.

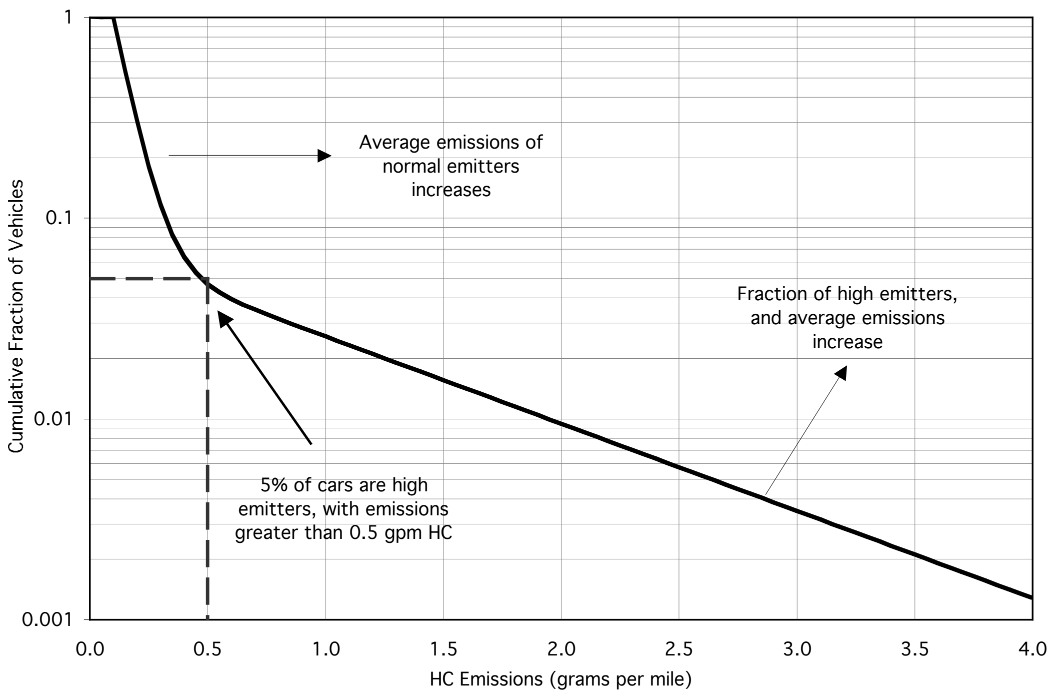


Fig. 2. Idealized VPD Trends with Mileage.

Two different processes account for the increase of gpm emissions over vehicle life:

> a) moderate and probably gradual degradation of emissions controls as vehicles age and accumulate mileage
> b) increasing fraction of vehicles with severely malfunctioning or failed controls

Degradation characterizes "normal emitters". It is a process affecting most vehicles, with moderate consequences. Unlike degradation of normal emitters, "high-emitters" arise probabilistically. High emitters tend to be associated with poor maintenance (both neglect of maintenance and incompetent repairs) and with fragile controls. Four major kinds of malfunction/failure occur: runs lean, runs rich, misfire and catalyst failure, have comparable probabilities in the in-use fleet, but different consequences (Wenzel & Ross 1998).

The incidence of high emitters can be defined in different ways. In this study relatively high cutpoints are used. Our cutpoints are the same for the different model years.

Background on IM240. Large IM240 surveys enable the study of emissions by cohorts of cars and light trucks depending on model year, odometer range, test site and date, etc. The data considered here are from Arizona, Colorado and Wisconsin, and for vehicles **with more than 100,000 miles**, in order to emphasize high emitters.

The IM240 is a 240-second dynamometer cycle with varying load, similar to that of bag 3 of the FTP; measurements are typically expressed as grams per mile over the test. Vehicles are supposed to start the test with engine hot.

Unfortunately for statistical applications of the data, many of the measurements are abbreviated or ended at less than 240 seconds according to "Fast-Pass" and sometimes "Fast-Fail" procedures. In most cases, only a small fraction of all measurements involve the full test; so in order to obtain good statistics abbreviated measurements are often "adjusted" to approximate full test values. For this study, the fast-fail procedure may significantly increase errors in the incidence of high emitters in the Arizona IM240 data. Colorado and Wisconsin do not use Fast Fail.

The Observed Decline in High Emitters (IM240). With our cutpoints, very roughly 5 times the certification standards met by the cars when new, the fraction of cars called high CO or HC emitters at over 100,000 miles is 1 to 3% for model years 1993-95 in Wisconsin and Arizona, but higher in Colorado. (**Table 1; see also Figs. 3 through 9**.) The incidence is 10 to 20% for model years 1987-89. These numbers are sensitive to the cutpoints we selected. We chose the cutpoints in part from the shape of the VPD for each pollutant, where the shape/slope of the VPD appears to change as illustrated in Fig. 2. This procedure cannot be applied to the NOx distribution.

Table 1. Incidence of High Emitters, MY93-95 (IM240)

state IM240 program	odometer range 1000 miles	test years	Cutpoints: CO,HC,NOx	CO	HC	NOx
Arizona	150-200	1996	15, 1.5, 2	3%	3%	10%
Colorado	100-200	1996-97	20, 2, 2.5	8%	5%	12%
Wisconsin	100-200	1996-97	15, 1.5, 2	2%	1%	6%

As seen in the IM240 data, the incidence of high emitters and their contributions to emissions inventories decline with model year throughout the period to 1995 (the most recent year with significant data at high mileage). **As shown in Table 2, the typical decline from MY88 to MY94 is about 80%, i.e. by a factor of one-fifth. This corresponds to an annual decline of almost 25% per year.**

Table 2. Percent Reductions in High-Emitter Contribution: MY94 Relative to MY88 (IM240 data – test years and cutpoints carry over from Table 1)

State IM240 program	odometer range 1000 miles	adjusted fast fail?	CO	HCs	NOx
Arizona	150-200	Y	80%	85%	~65%
Colorado	100-200	N	55%	65%	~30%
Wisconsin	100-200	N	80%	85%	~80%

Note: The high-emitter contribution depends on mileage but not on vehicle age as such. Using data from different test years enables a check of this hypothesis; and the hypothesis is confirmed.

Note: The same cutpoints shown in Table 1 are used for all MY groups.

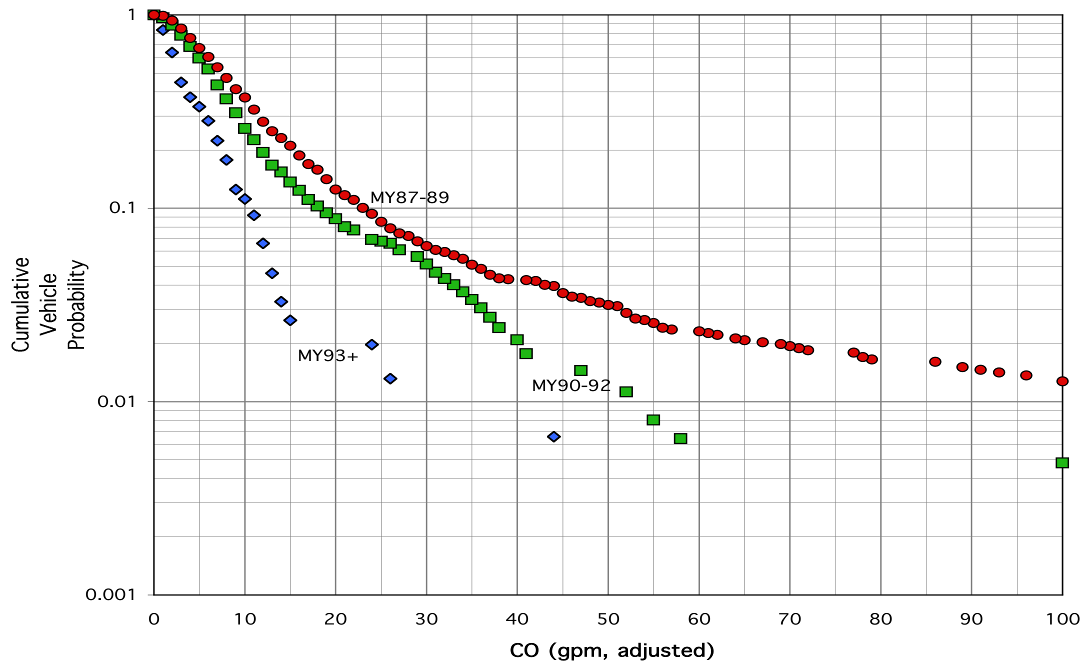


Fig. 3. CO Distribution, Arizona IM240, 1996 measurements by Model-Year Group; fuel-injected cars with 150,000 to 200,000 miles

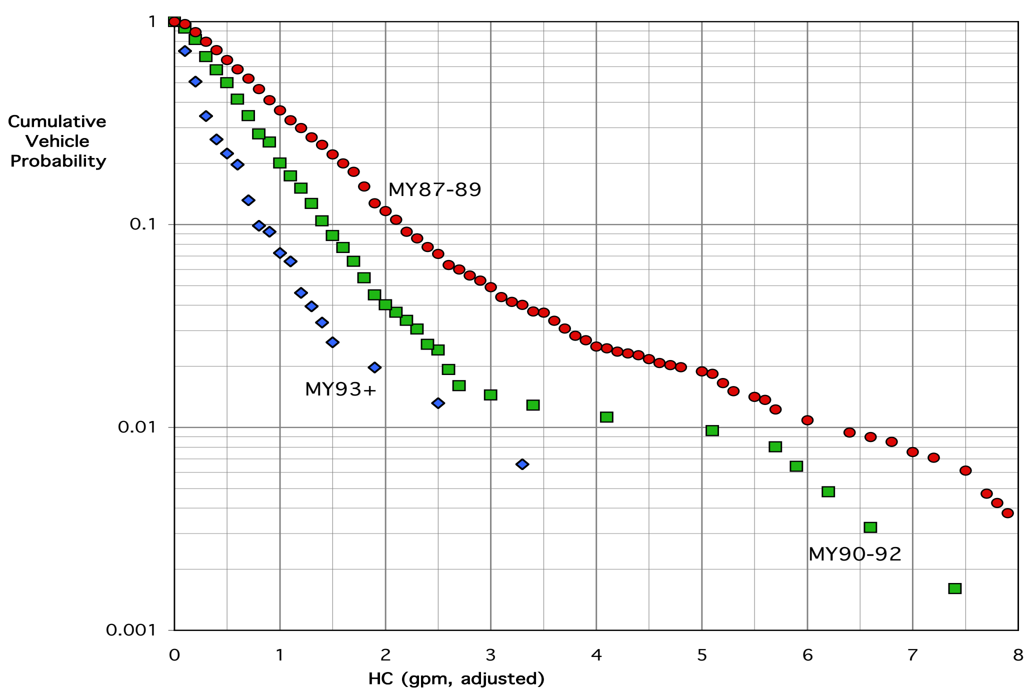


Fig. 4. HC Distribution, Arizona IM240. 1996 measurements by Model-Year Group: fuel-injected cars with 150,000 to 200,000 miles.

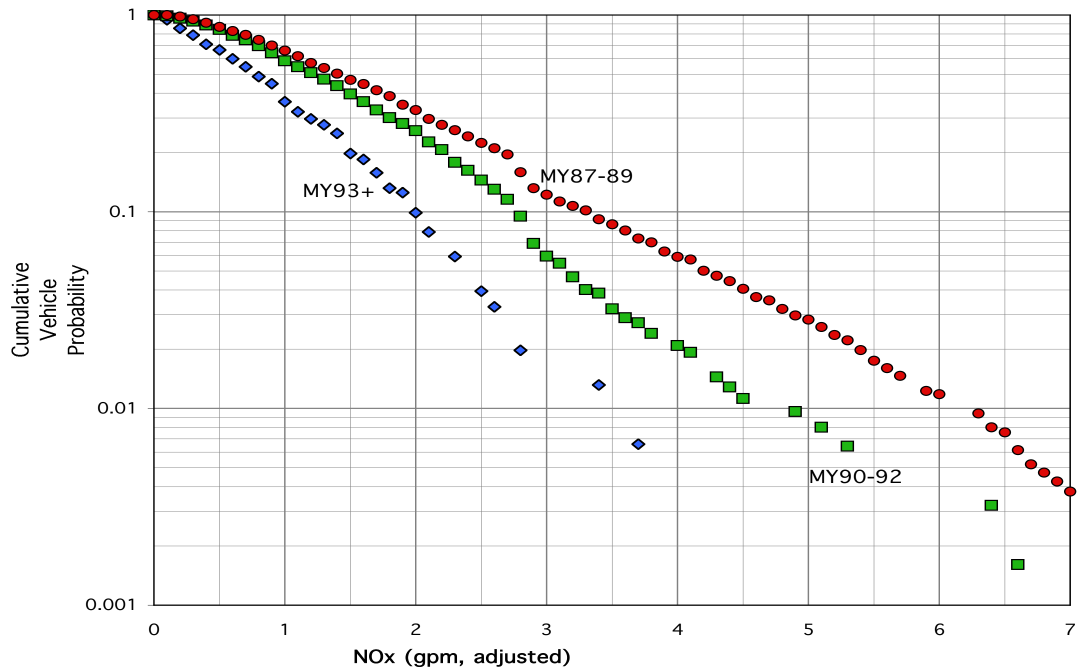


Fig. 5. NOx Distribution, Arizona IM240, 1996 measurements by Model-Year Group: fuel-injected cars with 150,000 to 200,000 miles.

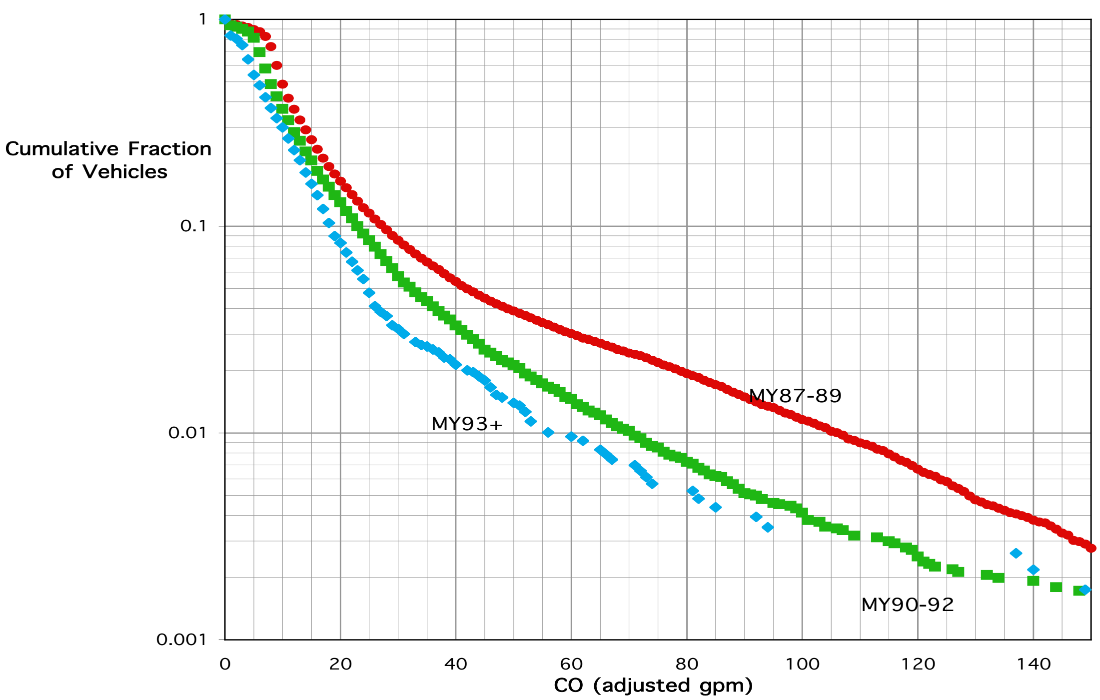


Fig. 6. CO Distribution, Colorado IM240, 1996-97 measurements by MY group; fuel-injected cars with 100,000 to 200,000 miles.

The declines are fairly consistent for HC and CO, which is expected because the enrichment and catalyst failure mechanisms are similar. (Although high HC emissions can also be caused by misfire.) They are also similar in Arizona and Wisconsin; but the progress observed in Colorado is less. We just show one of the Colorado VPDs (Fig. 6). For Colorado, for all three pollutants, not only is the progress with model year smaller than in Arizona and Wisconsin, but the emissions are significantly higher at a given cumulative fraction. We speculate that this is the result of damage to emissions controls in those vehicles that experience sustained high power driving in the mountains, rather than due to differences among the testing procedures.

A question that arises is whether or how strongly the results are influenced by prior I & M programs. In Arizona and Wisconsin the measurements are for the first IM240 test to which a vehicle has been subject. Vehicles may have been subject to simpler (idle test) programs. In Colorado, some of the vehicles went through the 1995 IM240 program prior to the measurements reported here.

The Wisconsin results are very similar to those from Arizona (Figs. 7 – 9) but more convincing because Fast-Fail is not used. We will focus on Arizona in the following, however, because of the good statistics and the availability of remote sensing data.

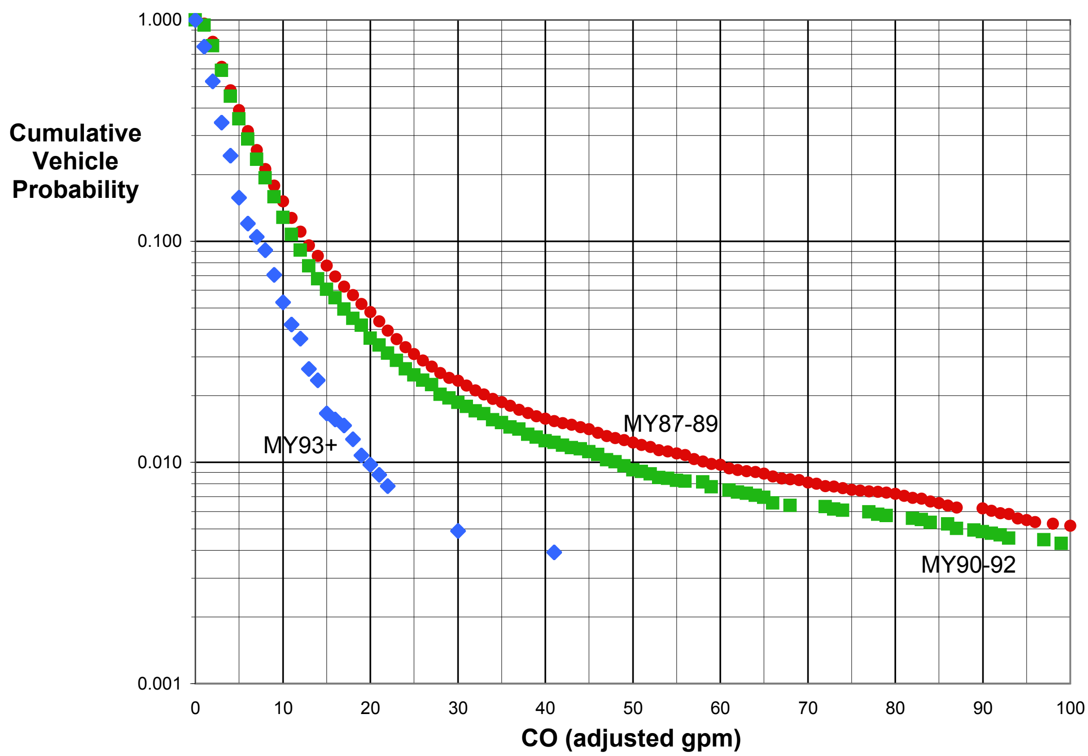


Fig. 7. CO Distribution, Wisconsin IM240, 1996-97 measurements by Model-Year Group: fuel-injected cars with 100,000 to 200,000 miles.

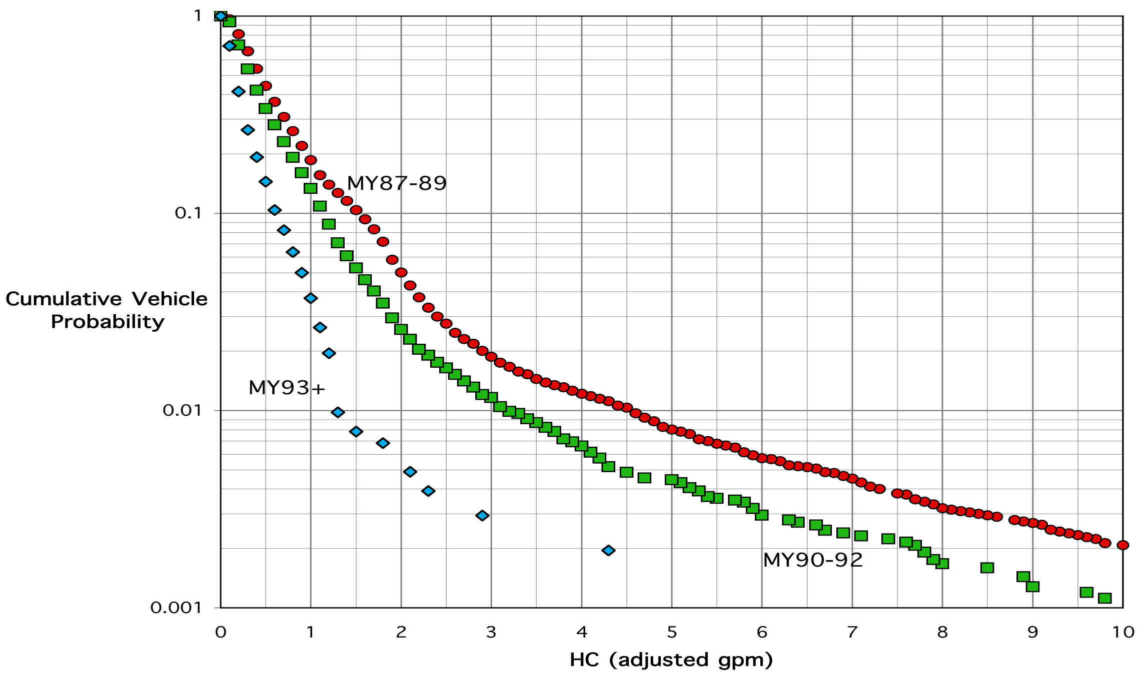


Fig. 8. HC Distribution, Wisconsin IM240, 1996-97 measurements by Model-Year Group; fuel-injected cars with 100,000 to 200,000 miles.

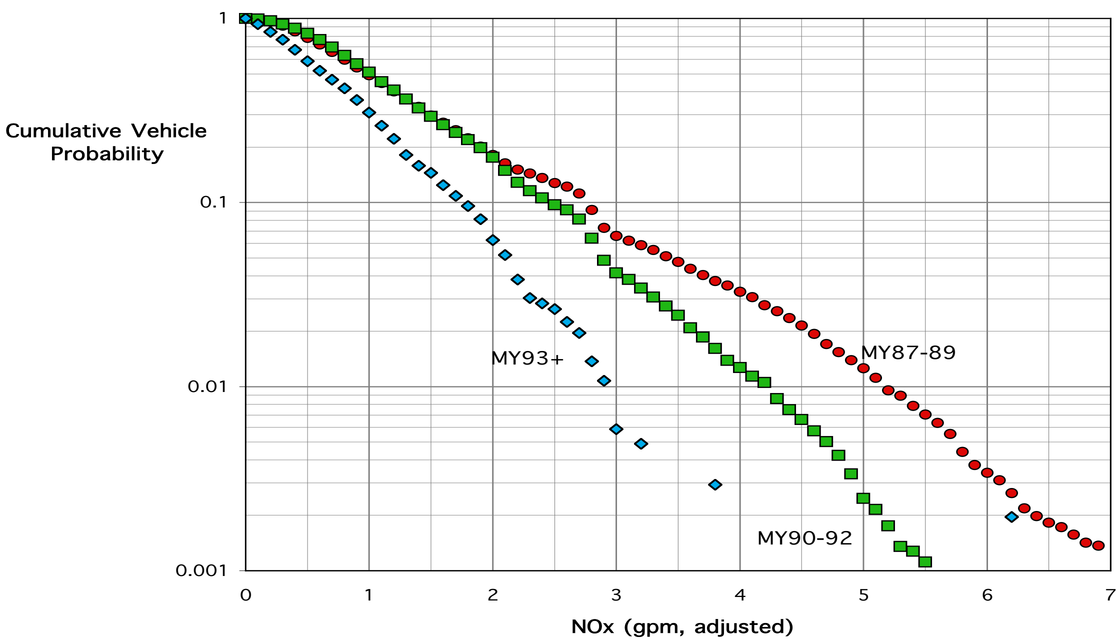


Fig. 9. NOx Distribution, Wisconsin IM240, 1996-97 measurements by Model-Year Group; fuel-injected cars with 100,000 to 200,000 miles.

Questions About Use of IM240 Surveys for Such Analyses. There are two major questions about the accuracy of statistical evidence on high emitters from IM240 surveys:

a) How accurate are the measurements?
b) Are there important biases in "recruitment" of vehicles?

Considerable inaccuracies are observed in IM240 measurements at low emissions levels. This is shown by comparing with FTP surveys for the same general cohort of cars. (See, e.g., **Fig. 10.**) In almost all cases the Arizona and Colorado IM240 distributions have too few vehicles with low gpm; i.e. the VPDs are too high at low gpm. Moreover, IM240 measurements by the Automotive Testing Lab have relatively more low gpm observations than the Arizona IM240 program, even though ATL used the same equipment. However, ATL used their own staff and procedures, suggesting that the most important part of the discrepancy arises from "poor conditioning", i.e. low engine/catalyst temperatures before testing in IM240 programs. This is strongly supported by back-to-back IM240 tests by Sierra Research, which found a substantial drop in emissions in the second test (Heirigs & Gordon 1996) However, the impact of poor conditioning on the analysis of high emitters is likely to be small, because the levels of cold start emissions in a normal emitter are smaller than the cutpoints used to define high emitters here.

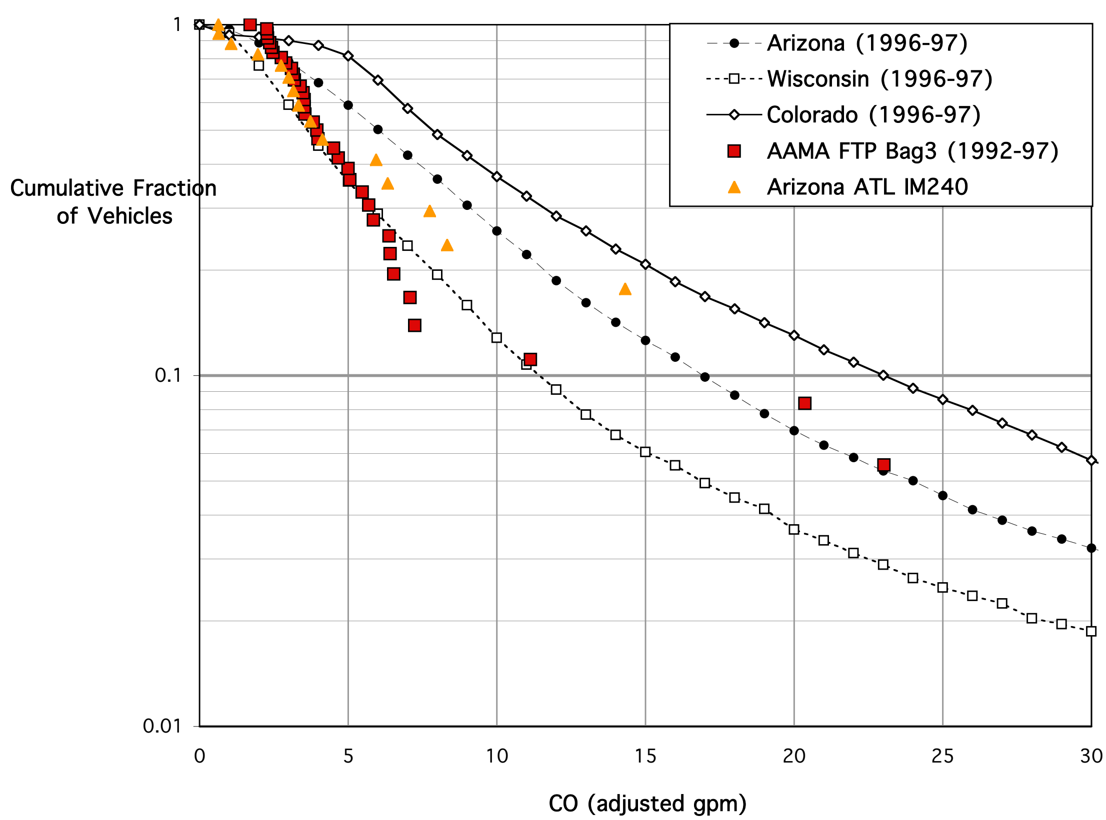


Fig. 10. CO Distributions for Model-Year 1990-92 Fuel-Injected Cars with 100 to 200k Miles. Comparison of AZ, CO and WI IM240, and ATL IM240 and AAMA FTP.

The behavior shown in Fig. 10 is for CO. A similar scarcity of vehicles with very low emissions is found for HCs and NOx, in the regular IM240 programs.

A second, more serious, caveat for us concerns "recruitment". Some cars and light trucks "required" to be tested simply aren't brought in to be tested; we do not have an estimate of that fraction. In the Phoenix area, of those vehicles brought in for testing in 1995, about 15% failed the initial test. Of these, some 40% never passed, although they are required to be brought back in after repairs until they do so. Of that group of "no-final-pass" vehicles, about half are still being driven in the Phoenix area more than 2 years after their 1995 I/M test, as shown by analysis based on different sets of vehicles seen at remote sensing sites (Wenzel 2000). One way people avoid the test is to move their registration outside of the IM240 region. Net registrations of roughly 10% of all vehicles due to be tested were moved outside an IM240 zone in Ohio when the enhanced test was introduced there (McClintock 1999). In addition some vehicles may not be tested because it appears risky to the testers (e.g. bald tires which would cause the vehicle to slip on the dynamometer, leaks in exhaust lines, etc). We conclude that since the fraction of avoiders is comparable to or larger than the number of high emitters, and since avoiders are likely to correlate strongly with high emitters, there is a serious statistical problem. The trend with model year may well be robust in the face of this difficulty; but any results from IM240 distributions need to be confirmed.

Background on Remote Sensing. Remote sensing surveys involve "snapshots" of concentrations of the pollutants at the tailpipe (ratios to CO2) as a vehicle passes a light beam across a single lane of traffic. A picture of the license plate is simultaneously taken. The snapshot is indeed brief; each cylinder having an exhaust stroke only about once per yard at 30 mph.

Vehicles passing remote sensing sites are not "recruited" in the sense they are for IM240. However some vehicles may avoid known sites. Moreover, more than half of the vehicles reporting for I/M testing in the Phoenix area were not observed in the extensive remote sensing program. So some kind of selection process is, in effect, going on. (The Phoenix remote sensing program was established to identify suspected high emitting vehicles, so there was some incentive for drivers to avoid detection by the sensors; but that may not be a major reason for the absence of so many vehicles.) Problems with the Phoenix remote sensing data have been noted elsewhere (Wenzel 1999).

Remote sensing sites may involve grades and certainly involve a distribution of speeds & accelerations. So the load is not usually known. Speed, acceleration and grade are being recorded in some recent surveys.

The Observed Decline in High Emitters (RSD). Although odometer readings are not determined directly in remote sensing, the approximate odometer reading can be inferred by matching the license plate with registration or I/M records of individual vehicles. The decline in high emitter contribution in RSD for the period MY88-MY95 is shown for CO & HC in **Figs.11 & 12**, for 100-200k odometer groups as observed in the Phoenix area. For cars, the decline in high emitter contribution in RSD for the period MY88-MY95 is seen to be roughly 65 to 70%% for both CO and HC, as compared with the 80 to 85% for the Arizona IM240 measurements (Table 2). **This is good agreement** considering the differences between the programs.

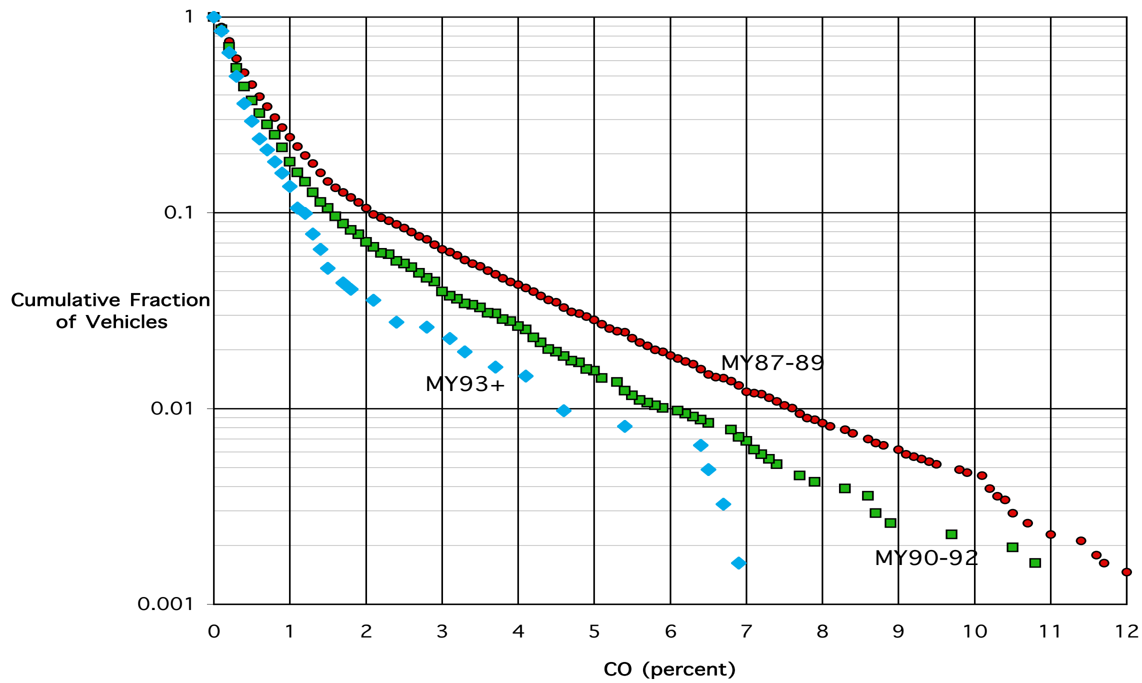


Fig. 11. CO Distribution, Arizona Remote Sensing, 1996-97 measurements by Model Year Group fuel-injected cars, with approximately 100,000 to 200,000 miles, measured up to three months before an IM240 test.

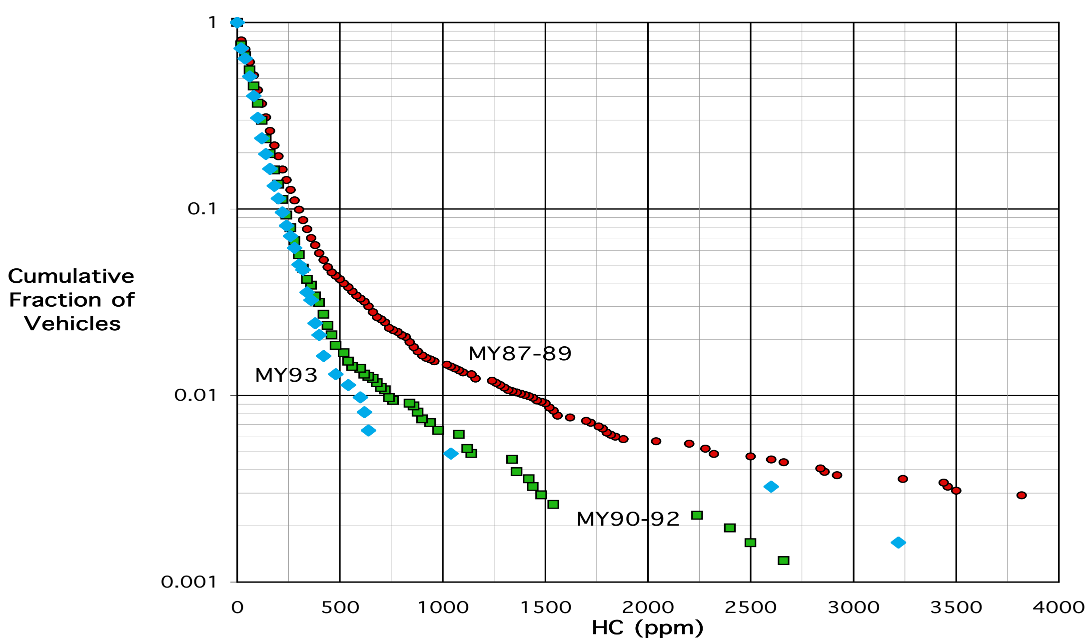


Fig. 12. HC Distribution, Arizona Remote Sensing, 1996-97 measurements by Model Year Group fuel-injected cars, with approximately 100,000 to 200,000 miles, measured up to three months before an IM240 test.

One feature of the IM240-RSD comparisons is that where there is a sharp decline in IM240 emissions from MY90-92 to MY93-95, there is a relatively small decline in the RSD. (We would like to thank Brett Singer for bringing attention to this feature.)

Comparison of IM240 and RSD. RSD concentrations are not directly comparable with gram/sec or gpm IM240 (or FTP) data. However, both IM240 and RSD include measurement of CO2 and so the pollution index, or mass ratio of pollutant to fuel used, can be determined in both cases. (We could also calculate concentrations using the IM240 data, but index is preferred because it reduces the variation associated with vehicle/engine size.)

For general background information, the correspondences between gpm, concentration and index are, for CO:

(conc in %) = 0.076*(index in %)

(index in %) $\approx CO_{gpm}/1.15$

For HC:

(conc in %) = 0.049*(index in %)

(index in %) $\approx HC_{gpm}/1.15$

where we relate the index % to the gpm using the fuel rate for MY1990s cars in the Arizona full IM240 sample (115 ± 25 gpm). Thus a CO concentration of 1% corresponds to a CO index of 13%, and emissions of roughly 15 gpm for CO. A HC concentration of 500 ppm, or 0.05%, corresponds a HC index of 1% and emissions of roughly 1.2 gpm for HC. A measure similar to the index is grams of pollutant per gallon of fuel, where gpg $\approx$ 2800*index. The density of gasoline varies significantly and the index (grams per gram) is more rigorous, so we use indices.

The pollution indices deduced from IM240 gpm measurements are:

$$FR_{gpm} = 13.9[CO_{gpm}/28 + CO2_{gpm}/44 + HC_{gpm}/13.9]$$

$$CO_{index} = CO_{gpm} / FR_{gpm}, \quad \text{and so on.}$$

Here FR stands for fuel rate and the range is the standard deviation.

The indices deduced from RSD concentrations are determined:

$$z \equiv CO\% / CO2\% = 6.56CO\% *(100 - 4.68CO\%)$$

$$M(\text{exhaust gases})/M(CO2) \equiv X = 1+ z +3.77(1+ z/2+ (1+z)n/4) = 6.56 + 4.68z$$

$$CO\% = [(12 + n)/ 28(1 + 3.77(1 + n/4))](CO_{index} \text{ in } \%) = 0.0757\ CO_{index}$$

$$HC_{index} = X*HC\% / [100(1 + z)]$$

Where M is moles; n is the H/C atom ratio in the fuel; and n =1.9 in the numerical expressions; and the carbon in HCs is neglected in calculating the total C in the exhaust.

CO indices as measured by both methods are shown in **Fig. 13** for CO. **The agreement between the two different methods of measurement and program design is surprisingly good**. As of this presentation, HC indices have not been calculated from the RSD; there are issues to be resolved (Singer et al 1998).

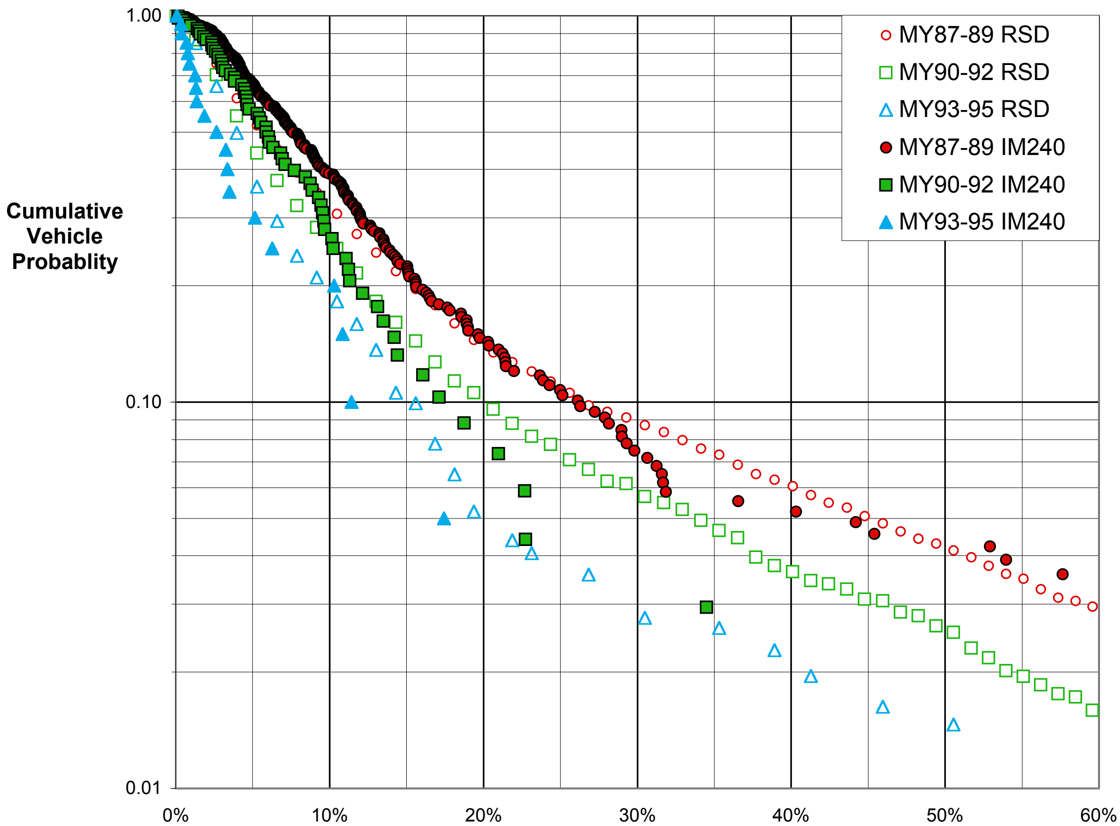


Fig. 13. CO Index Comparison: Arizona IM240 and Remote Sensing by Model-Year Group; 1996-97 measurements; fuel-injected cars, with approximately 100,000 to 200,000 miles.

It would be desirable to reduce one major uncertainty in this comparison, by reducing the difference in the power distributions (Jiminez & McClintock, CRC 1999). Unfortunately we cannot do this for the Arizona RSD. There are problems with the accuracy of the speed and acceleration measurements in the Phoenix RSD; in addition, some of the sites were on negative grades, and the grades of individual sites are not known.

Simulating Concentration Distributions Using Second-by-Second IM240. Something much simpler is to see how much the selection of high or low specific power seconds in the IM240 data changes the index or concentration distribution. (Index and concentration distributions have the same behavior.) In Fig. 14, distributions of the CO index at high-power (sec. 143) are compared with that at relatively low power (sec. 103) and for the whole IM240, for a dataset where second-by-second data is available. (The specific power, 2va, is 98 mph^2/s at sec. 103, and 5 at sec 143.) It is seen that the high-emitter emissions differ by a factor of two at the same percentile. So even over the range of the driving sampled by the IM240 cycle, emissions by high-emitters differ greatly.

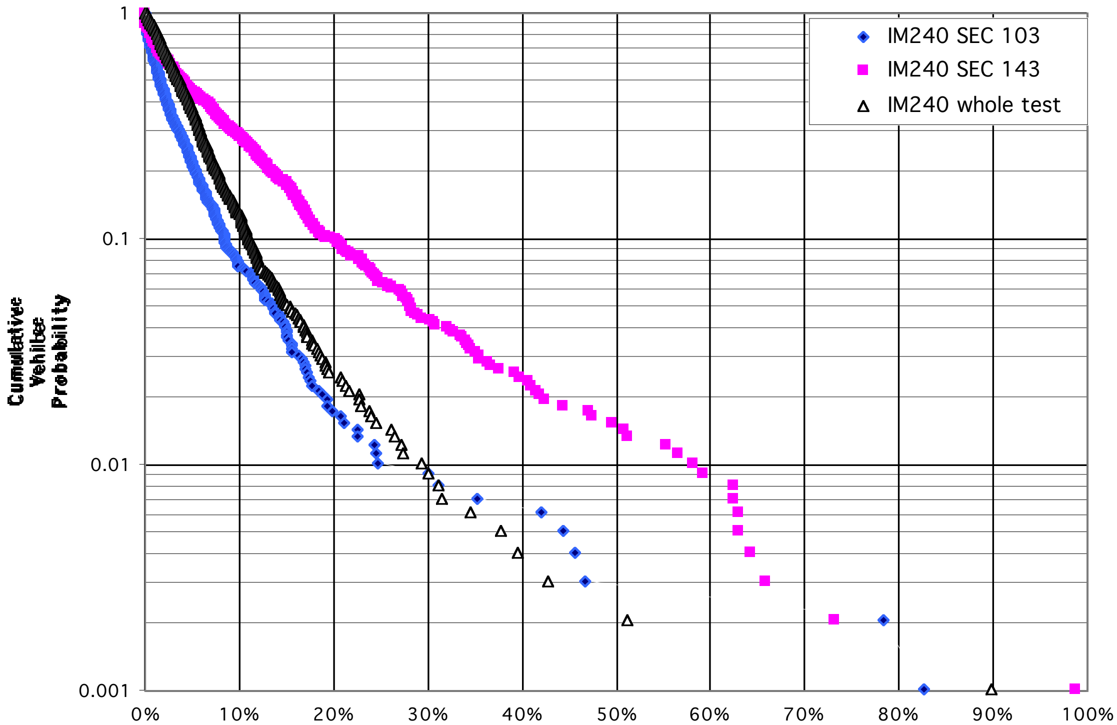


Fig. 14. CO Index distribution: Arizona IM240 at Different Power Levels. MY90-92 Cars of all mileages, "2% random" tests first half 1996.

From a scattergram of CO concentration and gpm based on the same IM240 second-by-second measurements (not included), one finds that at low-power points, the concentration distribution is spread out, with many high values. The proximate mechanism for the excess high concentrations is that while the emissions in grams are relatively low, the fuel rate/exhaust flow is very low. Thus observing a vehicle operating at low power by remote sensing can suggest incorrectly that the vehicle is a high emitter in gpm. This comparison is not pursued further.

CONCLUSIONS

• IM240 and remote sensing surveys agree in showing a striking decline with model year in the incidence of high emitters and their contribution to emissions inventories. For cohorts of cars with 100 to 200k miles, for model year 1994, the grams-per-mile emissions by high emitters were only 1/5 to 1/3 of those for MY 1988. We believe this success in reducing the emissions from vehicles with malfunctioning controls is associated with the increasing durability inherent in the control technologies developed and applied by manufacturers, in conjunction with systematic efforts of regulators to encourage improved certification and in-use emissions performance.

• IM240 and remote sensing measurements can be directly compared in terms of pollution indices; and they agree rather well, at least for CO from high-mileage cars. There is, however, a tendency for the emissions as measured by IM240 to decline sharply for the most recent model years studied (MY93-95), while the RSD show only a slight decline from MY90-92 to MY93-

95. This difference may arise because the two measurement programs sample somewhat different kinds of driving (such as cold vs hot engine/catalyst, or low vs high power). And changes in emissions behavior with those model years may focus on particular kinds of driving. For example, MY93-95 emissions may have been especially reduced in hot moderate-to-high power driving (more than emissions in other kinds of driving); and IM240 may mainly sample that kind of driving more strongly than does RSD. Then emissions as measured in IM240 might be reduced much more than those measured in remote sensing.

REFERENCES

Heirigs, Phillip L., and Jay Gordon. "Preconditioning Effects on I/M Test Results Using IM240 and ASM Procedures", Society of Automotive Engineers 962091, 1996.

Jimenez, Jose L., Peter McClintock, G.J. Mcrae, David D. Nelson, and Mark S. Zahniser, "Vehicle Specific Power: A Useful Parameter for Remote Sensing and Emission Studies," Proceedings of the 9th Coordinating Research Council On-Road Vehicle Emissions Workshop, San Diego, CA, April 21, 1999.

McClintock, Peter, I/M Impact on Dayton Registration Transactions, memo to EPA OMS, July 5, 1999.

Singer, Brett, Robert Harley, David Littlejohn. Jerry Ho & Thu Vo, "Scaling of Infrared Remote Sensor Hydrocarbon Measurements for Motor Vehicle Emission Inventory Calculations, **Environmental Science & Technology**, v 32, pp 3241-3248, 1998.

Wenzel, Tom, and Marc Ross, "Characterization of Recent-Model High-Emitting Automobiles", Society of Automotive Engineers 981414.

Wenzel, Tom, "Lessons Learned from I/M Program Evaluation", report to NRC Committee on Effectiveness of Vehicle I & M Programs, Feb. 2000.

Wenzel, Tom, "Evaluation of Arizona's Enhanced I/M Program," Proceedings of the 9th Coordinating Research Council On-Road Vehicle Emissions Workshop, San Diego, CA, April 21, 1999